Daniel C. Dennett

LA PELIGROSA IDEA DE DARWIN



A pesar de haber sido sometida al continuo acoso de quienes han tratado de limitar o negar sus implicaciones, la teoría darwiniana de la evolución por selección natural, una de las más controvertidas de todos los tiempos, se mantiene vigente y abierta a nuevas interpretaciones y desarrollos.

El autor describe con gran agudeza la teoría de Darwin y extiende sus argumentos hasta sus últimas y sorprendentes consecuencias.

El lector tiene en sus manos una de las más profundas reflexiones sobre la significación intelectual del darwinismo, un audaz, innovador y atractivo análisis sobre el efecto de renovación que la teoría de Darwin ha supuesto en nuestra actual visión del mundo.

Daniel Dennett

La peligrosa idea de Darwin

ePub r1.0
Titivillus 16.05.2021

Título original: *Darwin's Dangerous Idea*
Daniel Dennett, 1996
Traducción: Cristobal Pera Blanco-Morales

Editor digital: Titivillus
ePub base r2.1

Índice de contenido

A Van Quine,
maestro y amigo

Prefacio

Siempre me ha fascinado la teoría de Darwin sobre la evolución por selección natural, aunque a lo largo de los años he tropezado con una sorprendente variedad de pensadores que no pueden ocultar su malestar, expresado con diferentes matices, desde un indulgente escepticismo hasta una manifiesta hostilidad. He encontrado no sólo a gente lega en la materia y a pensadores religiosos, sino también a filósofos, psicólogos, físicos e incluso biólogos, que hubieran preferido que Darwin estuviese equivocado. Este libro explica por qué la idea de Darwin es tan poderosa, y por qué promete —no amenaza— situar nuestras visiones más apreciadas de la vida sobre unos nuevos fundamentos.

Sólo unas pocas palabras sobre el método seguido. Este libro trata ampliamente de ciencia pero no es, en sí mismo, un trabajo científico. La ciencia no se hace citando a autoridades —aunque sean muy elocuentes y eminentes— y evaluando sus argumentos. Sin embargo, los científicos, de manera correcta, persisten en mantener sus tesis en libros y ensayos populares y no tan populares, dando a conocer sus interpretaciones de los trabajos de laboratorio y de campo, y tratando de influir en sus colegas científicos. Cuando los cito, incluso con su carga retórica, yo hago lo que ellos: procuro ser convincente. No existe eso que se llama argumento de autoridad, pero ésta suele ser persuasiva, a veces de manera correcta y otras errónea. Trato de poner orden en todo esto, aunque no domino toda la ciencia que es relevante para las teorías que discuto, cosa que, por otra parte, les sucede también a los científicos (quizá con unas pocas excepciones de sabios en varias disciplinas). El trabajo multidisciplinario tiene sus riesgos. En los detalles de varios temas científicos he ido, espero, lo suficientemente lejos para permitir que el lector no informado comprenda exactamente de qué cuestiones se trata y por qué he optado por una determinada interpretación, facilitando además numerosas referencias.

Los nombres con fechas se refieren a referencias completas recogidas en la bibliografía que aparece al final del libro. Las notas al pie permiten aquella digresión que no todos los lectores apreciarían o requerirían[1t].

En este libro he tratado de hacer posible que el lector tenga acceso a la literatura científica que cito, para facilitar una visión unificada de la cuestión, junto con sugerencias acerca de la importancia o no de las controversias surgidas. En algunas de las disputas tomo partido claramente, mientras que en otras dejo el tema abierto aunque en un contexto que permita al lector entender cuáles son las cuestiones y —si le interesa— cómo han surgido. Espero que el lector consulte esta literatura, llena de admirables ideas. Algunos de los libros citados se encuentran entre los más difíciles que yo he leído. Pienso en los de Stuart Kauffman y Roger Penrose, por ejemplo, que son *tours de force* pedagógicos de materias extraordinariamente complicadas, los cuales pueden y deben ser leídos por todo aquel que desee conseguir una opinión instruida acerca de los importantes temas que plantean. Otros libros son de lectura menos exigente —claros, informativos, bien merecen un serio esfuerzo— y aún otros, no fáciles de leer, son una delicia; ejemplos magníficos de arte al servicio de la ciencia. Quien lea este libro probablemente habrá leído ya algunos de ellos, por lo que citarlos ahora conjuntamente es una recomendación suficiente: me refiero a los libros de Graham Cairns-Smith, Bill Calvin, Richard Dawkins, Jared Diamond, Manfred Eigen, Steve Gould, John Maynard Smith, Steve Pinker, Mark Ridley y Matt Ridley. Ningún área de la ciencia ha sido tan bien servida por sus 'escritores como la teoría de la evolución.

No se encuentran en este libro ese tipo de argumentos filosóficos de gran complicación técnica que muchos filósofos prefieren. Esto es así porque tengo un problema prioritario del que ocuparme. He aprendido que los argumentos, no importa cuál sea su solidez, caen a menudo en saco roto. Yo mismo he sido autor de argumentos que considero rigurosos e incontestables pero que, con frecuencia, no es que sean refutados o no discutidos, sino sencillamente desdeñados. No me quejo de injusticia, ya que todos debemos pasar por alto argumentos, incluidos aquellos que la historia nos dice que debíamos haber tomado seriamente. Más que nada, deseo desempeñar un papel más directo en el cambio de mentalidad de los que consideran que estas cuestiones pueden no tenerse en cuenta. Mi objetivo es que los pensadores que trabajan en otras disciplinas tomen en serio el pensamiento evolucionista, demostrándoles que lo han subestimado y que han escuchado a las falsas sirenas. Para conseguir este objetivo he de utilizar métodos más ingeniosos. Tengo que contar una historia. ¿No desea el lector cambiar de cambiar de opinión mediante una historia? Yo sé que el lector no quiere hacerlo con argumentos formales; que

incluso no quiere oír hablar de un argumento formal como conclusión, así es que voy a comenzar por donde tengo que comenzar.

La historia que voy a contar es, en su mayor parte, nueva, aunque elaborada a partir de pequeños fragmentos y piezas procedentes de una muy amplia serie de análisis que he escrito durante los últimos veinticinco años, en relación con varias controversias y dilemas. Algunas de estas piezas se han incorporado en su totalidad al libro, con mejoras, y otras tan sólo se mencionan de pasada. Tengo la esperanza de que la parte del iceberg que desvelo en este libro sea lo suficiente para informar e incluso persuadir al neófito y, al menos, retar a mis oponentes limpia y firmemente. He procurado navegar entre la Escila de una toma de postura demasiado suave y la Caribdis de un combate dialéctico que busca la pulverización minuciosa de una tesis, y cuando me deslizo hacia terreno controvertido, advierto que lo estoy haciendo, dando al lector referencias sobre la postura contraria.

Al comienzo de su maravilloso nuevo libro. *Metaphysical Myths, Mathematical Practices: The Ontology and Epistemology of the Exact Sciences* (Cambridge, Cambridge University Press, 1994), mi colega Jody Azzouni da las gracias al departamento de filosofía de la Tufts University «por facilitarle un ambiente casi perfecto en el cual podía *hacer* filosofía». Yo desearía secundar su agradecimiento y su valoración. En muchas universidades se estudia filosofía pero no se hace —«apreciación de la filosofía» podríamos llamar a esto—, y en otras muchas la investigación filosófica es una actividad arcana que se realiza fuera de la vista de los estudiantes de pregrado y de la mayoría de los postgraduados. En Tufts yo *hago* filosofía en el aula y entre nuestros colegas, y los resultados demuestran, en mi opinión, que la declaración de Azzouni es correcta. Tufts me ha facilitado excelentes estudiantes y colegas y un escenario ideal para trabajar con ellos. En años recientes he dirigido un seminario sobre Darwin y filosofía para la formación de pregrado, en él se han forjado la mayor parte de las ideas de este libro. El penúltimo borrador ha sido leído, criticado y pulido por un seminario particularmente exigente dedicado a estudiantes de postgrado y de pregrado, a los cuales les agradezco su ayuda: Karen Bailey, Pascal Buckley, John Cabral, Brian Cavoto, Tim Chambers, Shiraz Cupala, Jennifer Fox, Angela Giles, Patrick Hawley, Dien Ho, Mathew Kessler, Chris Lerner, Kristin McGuire, Michael Ridge, John Roberts, Lee Rosenberg, Stacey Schmidt, Rhett Smith, Laura Spiliatakou y Scott Tanona. El seminario se

enriqueció con numerosos visitantes: Marcel Kinsbourne, Bo Dahlbom, David Haig, Cynthia Schossberger, Jeff McConnell, David Stipp. También deseo dar las gracias a mis colegas, especialmente a Hugo Bedau, George Smith y Stephen White, por sus valiosas sugerencias. Mi gratitud especial a Alicia Smith, la secretaria del Centro de Estudios Cognitivos, cuya eficacia en la búsqueda de referencias, confirmación de datos, consecución de permisos, borradores, actualización, impresión y correspondencia, así como coordinación general del proyecto en su globalidad, ha puesto alas en mis pies.

También me he beneficiado de los detallados comentarios de aquellos que han leído la mayor parte o la totalidad de los borradores de los capítulos: Bo Dahlbom, Richard Dawkins, David Haig, Doug Hofstadter, Nick Humphrey, Ray Jackendoff, Philip Kitcher, Justin Leiber, Ernst Mavr, Jeff McConnell, Steve Pinker, Sue Stafford y Kim Sterelny. Como es habitual, ellos no son responsables de aquellos errores de cuya corrección no pudieron persuadirme. (¡Si alguien no puede escribir un buen libro sobre la evolución con la ayuda de este magnífico grupo de editores, lo mejor es que lo deje!).

Otros muchos han contestado a cuestiones cruciales y clarificado mi pensamiento en docenas de conversaciones: Ron Amundsen, Robert Axelrod, Jonathan Bennett, Robert Brandon, Madeline Caviness, Tim Clutton-Brock, Leda Cosmides, Helena Cronin, Arthur Danto, Mark De Voto, Marc Feldman, Murray Gell-Mann, Peter Godfrey-Smith, Steve Gould, Danny Hillis, John Holland, Alastair Houston, David Hoy, Bredo Johnsen, Stu Kauffman, Chris Langton, Dick Lewontin, John Maynard Smith, Jim Moore, Roger Penrose, Joanne Phillips, Robert Richards, Mark y Matt Ridley, Dick Schacht, Jeff Schank, Elliot Sober, John Tooby, Robert Trivers, Peter Van Inwagen, George Williams, David Sloan Wilson, Edward O. Wilson y Bill Wimsatt.

Mi agradecimiento a mi agente, John Brockman, por conducir el gran proyecto a través de numerosas dificultades, y ayudarme a ver cuáles eran los caminos a seguir para conseguir un libro mejor. Gracias también a Terry Zaroff, quien con su capacidad como editora del original evitó muchos olvidos e inconsistencias y clarificó y unificó la expresión de muchos puntos. Y a Ilavenil Subbiah, que dibujó las figuras, excepto las 10.3 y 10.4, que fueron diseñadas por Mark McConnell con un sistema de trabajo Hewlett-Packard Apollo, utilizando el programa I-dea.

Por último, y más importante, mi agradecimiento y amor a mi mujer Susan, por su consejo, cariño y apoyo.

Daniel Dennett

Septiembre de 1994

PRIMERA PARTE

Comenzando por el medio

Neurath ha comparado la ciencia con un barco, el cual, si hemos de reconstruirlo, deberemos hacerlo plancha por plancha, mientras permanecemos a bordo y a flote. El filósofo y el científico se encuentran en el mismo barco…

El análisis de cómo se construye una teoría, que es lo que pretendemos, debe hacerse comenzando por el medio. Nuestros principios conceptuales son objetos de tamaño mediano, situados a media distancia, y nuestra introducción a ellos y a cualesquiera otros se produjo a la mitad del camino en la evolución cultural del género humano. Si se asimila este hecho cultural seremos algo más conscientes de la distinción entre comunicación e invención, sustancia y estilo, señales y conceptualización, de lo que somos respecto a la distinción entre las proteínas y los hidratos de carbono en la materia de nuestra ingesta. Podemos distinguir retrospectivamente los componentes que han contribuido a la construcción de la teoría, como distinguimos las proteínas y los hidratos de carbono mientras subsisten como tales.

Willard van Orman Quine, *Word and Object*

CAPÍTULO I
Dime por qué

I. *¿No hay nada sagrado?*

Cuando era niño, solíamos cantar con frecuencia, fuese alrededor del fuego del campamento de verano, en el colegio y en la escuela de los domingos, o bien reunidos en casa junto al piano. Una de mis canciones favoritas era *Dime por qué*. (Para aquellos cuyos recuerdos personales no incluyan este pequeño tesoro, se facilita la música en el apéndice de este libro. La sencilla melodía y la fácil línea armónica son sorprendentemente bellas).

> Dime por qué brillan las estrellas,
> dime por qué las hiedras se retuercen,
> dime por qué el cielo es tan azul.
> Entonces yo re diré por qué te amo.
>
> Porque Dios hizo que las estrellas brillaran,
> porque Dios hizo que las hiedras se retorcieran,
> porque Dios hizo el cielo tan azul.
> Porque Dios te hizo a ti, por eso te amo.

Al oír esta declaración tan directa y tan sentimental todavía se me hace un nudo en la garganta; ¡es una visión de la vida tan dulce, tan inocente y tan reconfortante!

Y entonces llegó Darwin y nos aguó la fiesta. ¿Qué hizo Darwin? Éste es el tema de este libro. Desde que se publicó *El origen de las especies* en 1859, la idea fundamental de Charles Darwin ha provocado reacciones intensas que varían desde la condena feroz hasta la fidelidad extática y, a veces, casi el celo religioso. La teoría de Darwin se ha visto injuriada y tergiversada tanto por parte de amigos como de enemigos. Se han apropiado de ella de manera deshonesta, tomándola prestada para recubrir de respetabilidad científica espantosas doctrinas políticas y sociales. Ha sido colocada en la picota al ser caricaturizada por sus oponentes, algunos de los cuales tratan de ser competitivos en las escuelas de nuestros hijos con la «ciencia de la creación», un patético batiburrillo de pía pseudociencia[1].

Casi nadie es indiferente a Darwin y nadie debería serlo. La teoría de Darwin es una teoría científica, una gran teoría, pero no sólo eso. Los creacionistas que se oponen tan amargamente tienen razón en una cosa: la peligrosa idea de Darwin penetra mucho más profundamente en el entramado de nuestras creencias fundamentales de lo que muchos de sus refinados apologistas han admitido hasta ahora, incluso a sí mismos.

La dulce y simple visión de la canción, tomada literalmente, ya ha quedado atrás para muchos de nosotros, aunque la recordemos con afecto. El amable Dios que amorosamente nos ha creado (a todas las criaturas, grandes y pequeñas) y que, para nuestra delicia, ha esparcido por el cielo las brillantes estrellas, *ese* Dios es, como Papá Noel, un mito de la infancia, y no algo en lo que un adulto en su sano juicio y no desesperado pudiera realmente creer. Ese Dios debe convertirse en un símbolo de algo menos concreto o ser abandonado por completo.

No todos los científicos y filósofos son ateos y muchos de los que son creyentes declaran que su idea de Dios puede vivir, en pacífica coexistencia, con el entramado de las ideas de Darwin e incluso apoyarlas. Su Dios no es el antropomórfico Dios Todopoderoso sino un Dios todavía merecedor, a sus ojos, de adoración, capaz de dar consuelo y significado a sus vidas. Otros buscan fundamento para sus elevadas preocupaciones en filosofías completamente seculares, visiones del significado de la vida que les evite caer en la desesperación, sin otra ayuda de concepto alguno de Ser Supremo que la del propio universo. Algo es sagrado para estos pensadores, aunque no lo llamen Dios sino, quizá. Vida, Amor, Bondad, Inteligencia, Belleza o Humanidad. Lo que ambos grupos comparten, a pesar de las diferencias en sus creencias más profundas, es la convicción de que la vida tiene significado y de que la bondad vale la pena.

Pero ¿puede mantenerse frente al darwinismo esta actitud de admiración y este planteamiento, en *cualquiera* de sus versiones? Para comenzar, están aquellos que consideran que Darwin descubrió el peor de los pasteles: el nihilismo. Sostienen los que así, piensan que si Darwin estaba en lo cierto la consecuencia sería que nada puede ser sagrado. Para decirlo sin remilgos, nada tendría sentido. ¿Puede calificarse esta postura como una reacción excesiva? ¿Cuáles son exactamente las implicaciones de la idea de Darwin? Y en cualquier caso, ¿ha sido científicamente probada o es todavía «una simple teoría»?

Quizás el lector pueda pensar que sería útil establecer una distinción: hay partes de la idea de Darwin que realmente han sido demostradas más allá de

toda duda razonable y otras que son prolongaciones especulativas de aquellas partes científicamente indiscutibles. De esta manera —con suerte— los hechos con solidez científica no tendrían implicaciones llamativas sobre la religión o sobre la naturaleza humana, o sobre el significado de la vida, mientras que aquellas partes de la idea de Darwin que más han molestado podrían ser puestas en cuarentena al considerarlas prolongaciones muy controvertidas o meras interpretaciones de las partes científicamente indiscutibles. Todo esto resultaría reconfortante.

Pero ¡qué le vamos a hacer!, esto es justamente volver atrás a las andadas. Persisten duras controversias en torno a la teoría de la evolución, pero los que se sienten amenazados por el darwinismo no deben preocuparse por estas discusiones. La mayoría de las controversias —si no todas— se ocupan de cuestiones puramente científicas; cualquiera que sea la postura ganadora el resultado no anulará la idea básica de Darwin. Esta idea, tan firme como cualquier otra idea científica, tiene realmente implicaciones de largo alcance con respecto a lo que es, o pudiera ser, nuestra visión sobre el significado de la vida.

En el año 1543, Copérnico propuso que la Tierra no era el centro del universo sino que giraba alrededor del Sol. Fue necesario que transcurriera un siglo para que la idea fuera entendida en su totalidad, lo que resultó ser una gradual y prácticamente indolora transformación. (El reformador religioso Philipp Melanchthon, un colaborador de Martín Lutero, opinó que «algún príncipe cristiano» debería eliminar a ese loco, pero aparte de algunos otros exabruptos de este tipo, la mayoría de la gente no se sintió conmocionada por Copérnico). A la revolución copernicana le llegó el momento de ser escuchada «como un disparo que diera la vuelta al mundo» cuando Galileo Galilei escribió su *Dialogo sopra i due massimi sistemi del mondo tolemaico e copernicano*, pero no se publicó hasta 1632, en un momento en el que la idea ya no era motivo de controversia entre los científicos. El «proyectil» de Galileo provocó una infame respuesta de la Iglesia católica romana, poniendo en marcha una onda de choque cuyas reverberaciones no se están extinguiendo hasta ahora. Pero a pesar del drama de este enfrentamiento épico, la idea de que nuestro planeta no es el centro de la creación se ha asentado firmemente en la mente de la gente. Todos los niños en edad escolar aceptan hoy que esto es así, sin lágrimas ni terror.

Del mismo modo, y a su debido tiempo, la revolución darwiniana llegará a ocupar un lugar seguro y sin sobresaltos en las mentes —y en los corazones — de toda persona educada de nuestro planeta, pero hoy, transcurrido más de

un siglo de la muerte de Darwin, aún no hemos dado por terminada la discusión sobre aquellas implicaciones que inquietan a la mente. A diferencia de lo que ocurrió con la revolución copernicana, que no despertó la atención pública hasta que sus detalles científicos habían sido totalmente articulados, la revolución darwiniana ha tenido espectadores no expertos, llenos de ansiedad, y animadores que han tomado partido desde el comienzo, tirando de las mangas de los participantes y animando al público de la tribuna. Incluso los científicos se han visto afectados por idénticas esperanzas y temores, por lo que no debe sorprendernos que los conflictos relativamente escasos entre los teóricos hayan sido, a menudo, si no exagerados por sus partidarios, sí gravemente distorsionados durante el proceso. Todo el mundo ha entrevisto confusamente que hay muchas cosas en juego.

Por otro lado, aunque la articulación de la teoría de Darwin, realizada por él mismo, ha sido una obra monumental y su fuerza fue reconocida inmediatamente por muchos de los científicos y otros pensadores de su tiempo, existen en realidad amplios espacios vacíos en su teoría que sólo han comenzado a rellenarse recientemente. En una visión retrospectiva, algunos de los más amplios hiatos de la teoría parecen casi cómicos. En todas sus brillantes reflexiones Darwin nunca atinó con el concepto central, sin el cual poco cabe esperar de la teoría de la evolución: el concepto de *gen*. Darwin no contaba con la idea de una *unidad* de la herencia propiamente dicha y, por esta razón, su exposición del proceso de la selección natural está llena de dudas razonables acerca de si este proceso funcionaría. Darwin suponía que los hijos siempre exhibirían una especie de mezcla o término medio de las características de sus padres. ¿No sería siempre tal «mezcla heredada» un simple término medio que eliminaría las diferencias, haciendo que todas las cosas fuesen de un gris uniforme? ¿Cómo podía sobrevivir la diversidad con tal persistente tendencia al término medio? Darwin reconoció la importancia de este problema y ni él ni sus más ardientes defensores tuvieron éxito en el desarrollo de una respuesta que describiera un convincente y bien documentado mecanismo de la herencia que pudiera combinar los rasgos de los padres con el mantenimiento de una subyacente e inmodificable identidad. La idea que ellos necesitaban, y que era la correcta, la tenían a mano, descubierta (decir «formulada» sería demasiado fuerte) por el monje Gregor Mendel y publicada en una revista austríaca muy poco conocida en el año 1865, aunque en línea con las más sabrosas ironías de la historia de la ciencia, permaneció desconocida hasta que su importancia fue apreciada (en un principio de manera confusa) alrededor de 1900. Su triunfante

establecimiento como el corazón de la «síntesis moderna» —síntesis de Mendel y Darwin— se afirmó en la década de los cuarenta de este siglo, gracias a los trabajos de Theodosius Dobzhansky, Julian Huxley, Ernst Mayr y otros. Ha tenido que pasar medio siglo para conseguir alisar la mayoría de las irregularidades de este nuevo tejido de síntesis.

El núcleo fundamental del darwinismo contemporáneo, la teoría de la reproducción y la evolución basada en el ADN, está hoy más allá de cualquier disputa científica. Este núcleo fundamental demuestra su poder día a día, contribuyendo de manera crucial a que puedan explicarse desde los hechos a nivel planetario de la geología y la meteorología, pasando por los hechos a nivel medio de la ecología y la agronomía, hasta llegar a los hechos microscópicos de la ingeniería genética. Este núcleo fundamental unifica toda la biología y la historia de nuestro planeta en una única gran historia. Al igual que Gulliver inmovilizado en Lilliput, la teoría es inamovible no porque existan una o dos grandes cadenas de argumentos que podrían —esperanza contra esperanza— tener eslabones débiles, sino por estar firmemente amarrada por cientos de miles de cuerdas de evidencias ancladas virtualmente en todas las áreas del conocimiento humano. Es concebible que nuevos descubrimientos puedan conducir a cambios llamativos e incluso «revolucionarios» en la teoría de Darwin, pero la esperanza de que sea refutada por algunos progresos fulgurantes sería casi tan poco razonable como la esperanza de retornar a la visión geocéntrica y rechazar a Copérnico.

Y sin embargo, la teoría está envuelta en una ardiente controversia y una de las razones de este acaloramiento es que estos debates acerca de cuestiones científicas están habitualmente distorsionados por el miedo a que la respuesta «errónea» tenga intolerables implicaciones morales. Tan grandes son estos temores, que se mantienen cuidadosamente desarticulados, escamoteados de la atención por varias capas de rechazos y contrarrechazos. Los que disputan están continuamente cambiando, aunque sea ligeramente, el tema de la discusión, manteniendo a los demonios a buen recaudo. Este error de planteamiento es el principal responsable de que se demore el día en el que podamos vivir tan confortablemente con la nueva perspectiva biológica como lo hacemos con la perspectiva astronómica que Copérnico nos legó.

Siempre que se discute sobre el darwinismo la temperatura sube, porque lo que está en juego es algo más que hechos empíricos acerca de cómo ha evolucionado la vida sobre la Tierra o si es correcta la lógica que explica estos hechos. Una de las cosas más preciadas que se encuentran en peligro es una visión de lo que significa preguntar y responder, o sea, la pregunta «¿por

qué?». La nueva perspectiva introducida por Darwin pone patas arriba varias convicciones tradicionales, socavando los fundamentos de nuestras habituales ideas acerca de lo que deben considerarse como respuestas satisfactorias a esta antigua e inevitable cuestión. En este terreno la ciencia y la filosofía se encuentran completamente entrelazadas. Los científicos se engañan a sí mismos, a veces, cuando piensan que las ideas filosóficas son solamente algo decorativo o bien comentarios parásitos sobre los duros y objetivos triunfos de la ciencia, y que ellos mismos están inmunizados frente a las confusas cuestiones a cuya resolución dedican sus vidas. Pero no existe algo que pueda llamarse ciencia libre de la filosofía; sí existe en cambio la ciencia cuyo bagaje filosófico es asumido sin examen.

La revolución darwiniana es, al mismo tiempo, científica y filosófica, y ninguna de estas dos revoluciones podría haber tenido lugar en ausencia de la otra. Como más adelante veremos, fueron los prejuicios filosóficos de los científicos, más que la falta de evidencia científica, los que dificultaron que éstos se dieran cuenta de cómo la teoría podía realmente funcionar, pero aquellos prejuicios filosóficos que debían haberse eliminado estaban demasiado arraigados como para poder ser desalojados con un simple ejercicio de brillantez filosófica. Fue necesaria una secuencia irresistible de hechos científicos, conseguidos con muchas dificultades, para que los pensadores se vieran forzados a considerar seriamente la nueva y extraña visión de la vida propuesta por Darwin. Debe perdonarse la indeclinable lealtad a la ideas predarwinianas de aquellos que están aún mal informados con respecto a esa bella secuencia de hechos. La batalla no ha terminado todavía; incluso entre los científicos, quedan bolsas de resistencia.

Permítaseme poner las cartas sobre la mesa. Si se tratase de conceder un premio a la mejor idea que alguien haya tenido, yo se lo concedería a Darwin, por delante de Newton y de Einstein y algunos otros. De un solo golpe, la idea de la evolución por selección natural unifica la esfera de la vida, su significado y su propósito, con la esfera del espacio y el tiempo, de la causa y del efecto, de los mecanismos físicos y de las leyes que los rigen. Pero no se trata solamente de una admirable idea científica. Es, también, una idea peligrosa. Aunque mi admiración por la magnífica idea de Darwin no tiene límites, estoy de acuerdo, además, con muchas de las ideas y proyectos que *parecen* haber sido puestos en jaque por esta idea y deseo protegerlos. Por ejemplo, deseo proteger la canción del campamento, y la belleza y verdad que encierra, para mi pequeño nieto y para sus amigos y para sus hijos cuando nazcan y crezcan. Hay muchas más ideas magníficas que al parecer también

han sido puestas en difícil situación por la idea de Darwin, y que necesitan asimismo protección. El único camino adecuado para conseguirlo —la única vía que tiene una oportunidad a largo plazo— es atravesar las nubes de humo y examinar la idea de la manera más decidida e imparcial posible.

En esta ocasión, no vamos a decir aquello de «¡Allí, allí, todos juntos llegaremos!». Algunos pueden perder los nervios con nuestra indagación y sus sentimientos pueden ser heridos. Los que escriben sobre la evolución evitan este aparente conflicto entre ciencia y religión. Entrad deprisa necios, escribió Alexander Pope, allí donde los ángeles temen caminar. Lector, ¿quieres seguirme? ¿No deseas realmente conocer lo que ha sobrevivido de este enfrentamiento? ¿Qué pasaría si resultase que la dulce visión —u otra mejor— sobreviviese intacta, reforzada y profundizada tras el encuentro? ¿No sería una lástima haber perdido la oportunidad de conseguir un credo reforzado y renovado que sustituya a una fe frágil y enfermiza que se supone que no debe ser molestada?

No hay futuro en un mito sagrado. ¿Por qué no? A causa de nuestra curiosidad. Debido a que, como la canción nos recuerda, *deseamos conocer el porqué*. Podemos desechar la respuesta de la canción pero nunca podremos desechar la pregunta. Por muy apreciado que sea lo que poseemos, no podemos protegerlo de nuestra curiosidad, porque siendo lo que somos, la verdad es una de las cosas que más apreciamos. Nuestro amor por la verdad es seguramente un elemento central en el sentido que damos a nuestras vidas. En cualquier caso, la idea de que podemos conservar un significado engañándonos a nosotros mismos es más pesimista, más nihilista de lo que uno puede asimilar. Si esto fuera lo mejor que puede hacerse, yo concluiría que, después de todo, nada importa.

Este libro, en consecuencia, es para aquellos lectores que están de acuerdo en que la única explicación acerca de la significación de la vida de la cual vale la pena ocuparse es aquella que puede resistir con éxito nuestros mejores esfuerzos para examinarla. A los otros les aconsejo cerrar el libro y marcharse discretamente.

Para aquellos lectores que no abandonen el libro, he aquí mi plan. En la Parte I situaré la revolución darwiniana en el más amplio esquema de las cosas, mostrando cómo puede transformar la visión del mundo de aquellos que conozcan sus detalles. Este primer capítulo presenta el trasfondo de ideas filosóficas que dominaban nuestro pensamiento antes de Darwin. El capítulo 2 introduce la idea central de Darwin con un nuevo aspecto, considerando la idea de la evolución como un *proceso algorítmico* y aclarando algunos

malentendidos. El capítulo 3 muestra cómo la idea de Darwin desborda la tradición expuesta en el capítulo I. En los capítulos 4 y 5 se exploran algunas de las llamativas —e inestables— perspectivas que el darwinismo abre.

La Parte II examina los retos a la idea de Darwin —el neodarwinismo o «la síntesis moderna»— que han surgido dentro de la propia biología, mostrando que, contrariamente a lo que algunos de sus oponentes han declarado, la idea de Darwin ha sobrevivido a estas controversias y no sólo ha salido intacta sino reforzada. La Parte III muestra lo que ha sucedido cuando el mismo modo de pensar se extiende a la especie que nos interesa por encima de todas: el *Homo sapiens*. El propio Darwin reconoció que éste iba a ser el punto más hiriente para mucha gente e hizo lo que pudo para exponerlo con suavidad. Más de un siglo después, persisten aquellos cuyo deseo es abrir una zanja que nos separe de la mayor parte, si no de todas, de las temidas implicaciones que creen ver en el darwinismo. La Parte III muestra que esta actitud es un error de hecho y de estrategia; no sólo se aplica la peligrosa idea de Darwin a nosotros, los humanos, directamente y a muchos niveles, sino que la propia aplicación del pensamiento de Darwin a los problemas humanos —como son, por ejemplo, mente, lenguaje, conocimiento y ética— ilumina éstos de una forma que había sido siempre eludida en los planteamientos tradicionales, reconsiderando antiguos problemas y apuntando a su solución. Finalmente, podemos valorar el acuerdo conseguido cuando nos enfrentamos tanto con el pensamiento predarwiniano como con el darwiniano, identificando su uso y su abuso y mostrando que lo que realmente nos importa —lo que debe importarnos a todos— brilla a uno y otro lado, transformado aunque potenciado gracias a su travesía por la revolución darwiniana.

2. *¿Qué, dónde, cuándo, por qué y cómo?*

Tal como Aristóteles hizo notar en el nacimiento de la ciencia humana, nuestra curiosidad por las cosas se manifiesta de diferentes formas. Sus pioneros esfuerzos por conseguir una clasificación de estas formas están llenos de sentido. Identificó Aristóteles cuatro preguntas básicas para las que buscamos respuestas respecto a algo y llamó a sus respuestas las cuatro *aitia*, un término griego intraducible, aunque tradicionalmente traducido, de manera un poco extraña, por las cuatro ‘causas’.

1) Podemos mostrar curiosidad acerca de lo que algo está hecho, su materia o *causa material*.

2) Podemos mostrar curiosidad acerca de la forma (o estructura) que tal materia ha tomado, su *causa formal*.

3) Podemos mostrar curiosidad acerca de sus comienzos, es decir, cómo se inició, su *causa eficiente*.

4) Podemos mostrar curiosidad acerca de su *propósito, objetivo* o *fin* (como en la expresión «¿El fin justifica los medios?»), lo que Aristóteles llamó *telos*, a veces traducido, también extrañamente, como 'causa final'.

Es necesario un gran esfuerzo de comprensión y adaptación para conseguir que estas cuatro *aitias* de Aristóteles se acomoden como respuestas a las cuatro interrogaciones habituales «qué», «dónde», «cuándo» y «por qué». Esta acomodación es sólo parcialmente buena. No obstante, las preguntas que comienzan con «por qué» se corresponden habitualmenre con la interrogación por la cuarta «causa» de Aristóteles, el telos de una cosa. ¿Por qué esto?, preguntamos. ¿Para qué sirve esto? De la misma forma que se dice en francés ¿cuál es la *raison d'être*, o sea, la razón de ser? Durante siglos estos «porqués» han sido reconocidos como problemáticos por los filósofos y científicos, y tan distintos que los problemas que plantean merecen un nombre: teleología.

Una explicación *teleológica* es aquella que explica la existencia u ocurrencia de algo citando como prueba el objetivo o propósito al que sirve esa cosa. Los artefactos son los casos más obvios; el objetivo o propósito de un artefacto es la función a cumplir para la cual fue diseñado por su creador. No existe controversia acerca del *telos* de un martillo: golpear e introducir clavos. El *telos* de artefactos más complicados, como una cámara de vídeo con su pantalla de televisión portátil, un camión-grúa o un escáner para tomografía axial computarizada (TAC) es, incluso, más obvio. Pero aun en casos simples puede vislumbrarse confusamente en el trasfondo un problema:

> —¿Por qué estás serrando esta tabla?
> —Para hacer una puerta.
> —¿Y para qué es la puerta?
> —Para la seguridad de mi casa.
> —¿Y por qué deseas una casa segura?
> —Para poder dormir por las noches.
> —¿Y por qué deseas dormir por las noches?
> —Márchate y deja de hacer preguntas estúpidas.

Este intercambio de preguntas y respuestas revela uno de los problemas de la teleología: ¿dónde dejar de hacer preguntas? ¿Qué causa *final* puede ser invocada de modo que cierre esta secuencia jerárquica de razones? Aristóteles tenía una respuesta: Dios, el primer motor, con *quien* termina toda la serie de

interrogaciones acerca de *quién* lo ha hecho. La idea, asumida por las tradiciones cristiana, judía e islámica, es que todos *nuestros* propósitos son, en último término, propósitos de Dios. La idea es ciertamente natural y atractiva. Si observamos un reloj de bolsillo y nos preguntamos por qué tiene un cristal transparente en una de sus caras, la respuesta nos induce a pensar en las necesidades y deseos de quienes utilizan estos relojes, porque desean saber qué hora es, mirando a través del cristal transparente y protector, y así en adelante. Si no fuera por estos hechos que se relacionan con *nosotros*, para quienes el reloj fue creado, no habría explicación para el «porqué» de su cristal. Si el universo fue creado por Dios, para los propósitos de Dios, entonces todos los propósitos que pudiéramos encontrar en él deben ser debidos, en último término, a los propósitos de Dios. Pero ¿cuáles son estos propósitos? Es algo así como un misterio.

El desasosiego que produce este misterio puede atenuarse si cambiamos ligeramente de tema. En lugar de responder a la interrogación «¿por qué?» con una respuesta que comienza diciendo «porque» (el tipo de respuesta que parece requerir la pregunta), a menudo la gente sustituye la pregunta «¿por qué?» por la pregunta «¿cómo?» e intenta contestar contando una historia acerca de *cómo ha llegado a ser* que Dios nos creara, a nosotros y al resto del universo, sin pararse demasiado a considerar por qué Dios pudo desear hacer tal cosa. La pregunta «¿cómo?» no se encuentra en la lista de Aristóteles, pero era popular como pregunta y respuesta, antes de que éste realizara su análisis. Las respuestas a los «¿cómo?» como preguntas mayores son *cosmogonías*, historias acerca del *cosmos*, la globalidad del universo y todos los habitantes que han llegado a existir. El libro del Génesis es una cosmogonía, pero existen otras muchas. Los cosmólogos que exploran la hipótesis del Big Bang y especulan sobre los agujeros negros y las supercuerdas, son los modernos creadores de cosmogonías. No todas las cosmogonías antiguas siguen el modelo del hacedor de artefactos. Algunas implican a un «huevo del mundo» depositado en «lo profundo» por un mítico pájaro u otro ser y otras cosmogonías implican «semillas» sembradas y cultivadas. La imaginación humana posee escasos recursos para poner orden en estas enrevesadas preguntas. Un mito de la creación primitiva habla de un «Señor auto-existente» quien «con un pensamiento creó las aguas y depositó en ellas una semilla que se convirtió en un huevo de oro, del cual él mismo nació como Brahma, el progenitor de los mundos» (Muir 1972:IV, 26).

¿Y cuál es el significado de todos estos huevos depositados y todas estas semillas sembradas para la construcción del mundo? O ¿cuál es el significado

del Big Bang? Los cosmólogos actuales, como muchos de sus predecesores a través de la historia, relatan una divertida historia, aunque prefieren dejar a un lado el «porqué» que suscita la cuestión teleológica. ¿Existe el universo por alguna razón? ¿Puede la razón desempeñar algún papel inteligible en la explicación del cosmos? ¿Puede alguien existir por una razón, sin ser la razón de *alguien*? O ¿son razones —del tipo de las cuatro causas de Aristóteles— apropiadas solamente para la explicación de los trabajos y las acciones de los seres humanos u otros agentes racionales? Si Dios no es una persona, un agente racional, un Artífice Inteligente, ¿qué posible sentido puede tener el mayor «por qué»? Y si el mayor «porqué» no tiene sentido, ¿cómo podrían tenerlo otros «porqués» menores, de horizontes más limitados?

Una de las contribuciones más fundamentales de Darwin ha sido habernos mostrado una nueva vía por la que conseguir que los «porqués» tengan sentido. Guste o no guste, la idea de Darwin ofrece una vía —una vía despejada, convincente y sorprendentemente versátil— para solucionar estos viejos enredos. Cuesta tiempo aprender a utilizarla y a veces es incorrectamente recorrida, incluso por sus más fieles amigos. Exponer gradualmente y clarificar esta forma de pensar es un proyecto central del presente libro. El pensamiento darwiniano debe distinguirse cuidadosamente del pensamiento excesivamente simplificado y populachero de algunos impostores, lo que nos obligará a introducirnos en aspectos técnicos, pero merece la pena. El premio es, por primera vez, un sistema de explicación estable, que no gira una y otra vez, en círculos o en espiral, dando marcha atrás ante los misterios. Algunos preferirían, al parecer, el continuo retroceso ante los misterios, pero en estos tiempos el coste de esta actitud es excesivo; nos engañaríamos a nosotros mismos. El lector puede escoger entre engañarse a sí mismo o dejar a los otros el trabajo sucio, aunque no es una vía intelectualmente defendible volver a reconstruir las poderosas barreras que impedían la comprensión, aquellas que Darwin aplastó.

El primer paso para apreciar este aspecto de la contribución de Darwin es comprobar cómo era la visión del mundo, antes de que fuera trastocada por él. Si lo hacemos a través de los ojos de dos de sus paisanos, John Locke y David Hume, podremos conseguir una clara idea de una visión alternativa del mundo que aún persiste entre nosotros en muchos aspectos, visión que Darwin convirtió en obsoleta.

3. *La «prueba» de Locke sobre la primacía de la mente*

¡John Locke inventó el sentido común y desde entonces solamente lo poseen los ingleses!

Bertrand Russell [2]

John Locke, un contemporáneo del «incomparable señor Newton», fue uno de los padres fundadores del empirismo británico y, de acuerdo con su postura empírica, no era muy dado a los argumentos deductivos de los racionalistas, aunque una de sus nada características incursiones en el terreno de las «pruebas» merece ser citada en su totalidad, ya que ilustra perfectamente el bloqueo de la imaginación que existía antes de la revolución darwiniana. El argumento puede parecer extraño y pomposo para la mentalidad moderna aunque, al mismo tiempo, se puede considerar una muestra del largo camino recorrido desde aquellos tiempos. El propio Locke pensaba que lo que hacía era justamente recordar algo obvio. En este pasaje de su *Ensayo sobre el entendimiento humano* (IV, x, 10) el objetivo de Locke era *demostrar* algo en lo que él consideraba que todo el mundo creía seriamente: que «en el principio» era la Mente. Comenzaba preguntándose si algo era eterno:

Entonces, si algo debe ser eterno ¿qué tipo de Ser debe ser? Para la razón es muy obvio que debe ser necesariamente un ser cogitativo. Porque es imposible concebir que cualquier simple materia no cogitativa pueda producir un ser inteligente y pensante, así como que la nada de sí misma produzca materia…

Locke comienza su prueba aludiendo a una de las más antiguas y más utilizadas máximas filosóficas, *Ex nihilo nihil fit*: nada puede surgir de la nada. Dado que se trata de un argumento deductivo, debe apuntar alto: no sólo es improbable o no plausible o difícil de imaginar, sino *imposible de concebir* que «la simple materia no cogitativa pueda producir un ser inteligente y pensante». El argumento se desarrolla a lo largo de una serie de pasos ascendentes:

Supongamos que podemos encontrar una porción de materia eterna, grande o pequeña, en sí misma capaz de producir nada… La materia, por su propia fuerza no puede producir en sí misma movimiento: este movimiento debe ser también desde la eternidad o producido y añadido a la materia por algún otro ser más poderoso que la materia… Pero supongamos que el movimiento es también eterno; la materia, la materia no cogitativa y el movimiento, cualesquiera que sean los cambios que se puedan producir en su figura y en su masa, nunca puede producir pensamiento: el conocimiento queda fuera de las posibilidades de poder ser producido por el movimiento y la materia, del mismo modo que la materia no puede ser producida por la nada o el no ser. Y yo apelo a los propios pensamientos de cada uno, si él no puede fácilmente concebir que la materia sea producida por la nada, como el pensamiento no puede ser producido por la pura materia, cuando antes de ambos no existe tal cosa como el pensamiento o un ser inteligente…

Es interesante subrayar que Locke decide que él puede seguramente «dirigirse a los propios pensamientos de cada uno» para asegurar su «conclusión». Locke estaba seguro de que su «sentido común» era verdaderamente sentido *común*. ¿No consideramos como algo obvio que mientras que la materia y el movimiento pueden producir cambios de «figura» y de «masa» no puedan producir *nunca* «pensamiento»? ¿No se eliminaría de este modo la posibilidad de los robots, o al menos de los robots que pudieran generar genuinos pensamientos entre los movimientos producidos en los materiales de sus cabezas? Es cierto que en la época de Locke —que era también la de Descartes— la simple idea de una inteligencia artificial estaba tan cercana a lo impensable que Locke podía esperar con toda confianza un respaldo unánime a esta llamada a su audiencia, una llamada que hoy correría el riesgo de provocar carcajadas[3]. Como veremos más adelante, el desarrollo de la Inteligencia Artificial (*AI*) se deriva directamente de la idea de Darwin. Su nacimiento, que fue profetizado por el propio Darwin, vino acompañado por una de las primeras demostraciones impresionantes del poder formal de la selección natural (el legendario programa de damas de Art Samuel, que describiremos con algún detalle más adelante). Ambas, la evolución y la Inteligencia Artificial, inspiran el mismo disgusto en mucha gente que deberían conocerlas mejor, como veremos en próximos capítulos. Pero volviendo a la conclusión de Locke:

Así, si suponemos que nada es primero o eterno: la materia no puede nunca comenzar a ser. Si suponemos eterna la simple materia, sin movimiento, el movimiento no puede nunca llegar a ser; si suponemos primeros o eternos sólo la materia con el movimiento: el pensamiento no puede nunca llegar a ser. Por esto es imposible concebir que la materia, con o sin movimiento, pueda tener originalmente, en y por sí misma, sentido, percepción y conocimiento, como se deduce evidentemente, ya que entonces el sentido, la percepción y el conocimiento debían ser una propiedad eternamente inseparable de la materia y de todas sus partículas.

Por consiguiente, si Locke está en lo cierto, la mente debe ser lo primero, o al menos ligada a lo primero. No puede llegar a existir un momento después, como resultado de una confluencia de fenómenos más modestos, no inteligentes. Esto supone una vindicación totalmente secular, lógica —se podría decir que casi matemática— de un aspecto central de la cosmogonía judeocristiana (y también islámica): en el comienzo fue algo con mente —un Ser cogitativo—, como Locke dice. A la idea tradicional de que Dios es racional, agente pensante, un diseñador y constructor del mundo, se le concede aquí la más alta aprobación científica: como un teorema matemático, su denegación es aparentemente imposible de concebir.

Así les parecía a muchos brillantes y escépticos pensadores antes de Darwin. Casi cien años después de Locke, otro gran empírico británico, David Hume, se enfrentó de nuevo con el problema en una de las obras maestras de la filosofía occidental, sus *Diálogos sobre la religión natural* (1776).

4. *Hume estuvo muy cerca*

La religión natural, en tiempos de Hume, era una religión fundamentada en las ciencias naturales, en contraposición a la religión «revelada», que depende de una revelación (sea una experiencia mística o alguna otra fuente de convicción que no puede ser comprobada). Si el único fundamento para vuestras creencias religiosas es «Dios me lo dijo en un sueño», vuestra religión no es una religión natural. La distinción no tenía bastante sentido antes del nacimiento de la ciencia moderna en el siglo XVII, cuando ésta estableció un nuevo y competitivo criterio de evidencia para todas las creencias. De este modo se planteó la siguiente interrogación: «¿Puede facilitarnos algunos fundamentos *científicos* de sus creencias religiosas?».

Muchos pensadores religiosos, comprendiendo que el prestigio del pensamiento científico —en igualdad de condiciones— era una aspiración que merecía la pena, aceptaron el reto. Es difícil comprender por qué todo el mundo desea evitar la confirmación científica de su credo, si éste la tuviera. Entre los argumentos pretendidamente científicos con conclusiones religiosas, el favorito de manera abrumadora es el argumento del diseño en algunas de sus versiones. Entre los efectos que podemos observar objetivamente en el mundo, hay muchos que no son meros accidentes (no pueden serlo por varias razones); deben haber sido diseñados para ser como son y no pueden haber sido diseñados más que por un diseñador; en consecuencia, un diseñador, Dios, debe existir (o ha existido) como la fuente de todos estos admirables efectos.

Este argumento puede parecer un intento de buscar y encontrar una ruta alternativa a la conclusión de Locke, una ruta que nos llevaría a través de detalles más empíricos, en lugar de confiarnos tan franca y directamente en lo que se considera inconcebible. Las realidades de los diseños observados en el mundo pueden ser analizados, por ejemplo, para asegurar los fundamentos de nuestra apreciación de la sabiduría del diseñador y nuestra convicción de que el simple azar no puede ser responsable de estas maravillas.

En los *Diálogos* de Hume, tres personajes de ficción mantienen un debate con consumado ingenio y vigor. Cleantes defiende el argumento del diseño y lo hace con uno de sus más elocuentes discursos[4]. He aquí las palabras que lo inician:

Mirad el mundo que nos rodea: contempladlo en su totalidad y en cada una de sus partes. Encontrareis que no es otra cosa que una gran máquina, subdividida en un número infinito de máquinas más pequeñas, las cuales admiten subdivisiones hasta un grado que va más allá de lo que los sentidos y facultades humanas pueden seguir y explicar. Todas estas diversas máquinas, y hasta sus más diminutas partes, están ajustadas unas a otras con una precisión que provoca la admiración de todos los hombres que alguna vez las han contemplado. La curiosa adaptación de los medios a los fines a través de toda la naturaleza se asemeja exactamente, aunque la excede bastante, a los productos de la invención, del diseño del pensamiento, de la sabiduría y de la inteligencia humanos. Puesto que los efectos se asemejan unos a otros, estamos obligados a inferir, por todas las reglas de la analogía, que las causas también se asemejan y que el Autor de la naturaleza es algo similar a la mente del hombre, aunque poseedor de muchas más amplias facultades, proporcionadas a la grandeza del trabajo que ha ejecutado. Por este argumento *a posteriori*, y solamente por este argumento, nosotros probamos al mismo tiempo la existencia de Dios y su similitud con la mente y la inteligencia humanas (parte II).

Philo, el escéptico contrincante de Cleantes, desmenuza el argumento, preparándolo para ser destruido. Anticipándose al famoso ejemplo de Paley, Philo dice: «Arroja varias piezas de acero en un montón, sin forma alguna; nunca se dispondrán por sí mismas de modo que compongan un reloj»[5].

Prosigue Philo: «La piedra, el mortero y la madera, sin un arquitecto, nunca levantan una casa. Pero nosotros vemos que las ideas en una mente humana, por una desconocida e inexplicable economía, se disponen a sí mismas para formar el plan de un reloj o de una casa. Por lo tanto, la experiencia prueba que hay un principio original de orden en la mente, no en la materia» (parte II).

Debe señalarse que el argumento del diseño depende de una inferencia inductiva: donde hay humo hay fuego; y donde hay un diseño, hay una mente. Pero ésta es una inferencia dudosa, observa Philo: la inteligencia humana es

no más que una de las fuentes y principios del universo, como el calor o el frío, la atracción y la repulsión y cientos de otras fuentes, que caen dentro del ámbito de la observación cotidiana… Pero ¿puede una conclusión, con cualquiera propiedad, ser transferida de las partes al todo?… Observando el crecimiento de un cabello, ¿podemos aprender algo concerniente a la generación de un hombre?… ¿Qué privilegio peculiar tiene esta leve agitación del cerebro que nosotros llamamos pensamiento, para ser convertida por nosotros de este modo en el modelo de la totalidad del universo? ¡Admirable conclusión! La piedra, la madera, el ladrillo, el hierro y el cobre no tienen, en este tiempo, en este diminuto globo terráqueo, un orden o disposición sin el arte y la invención humana. Por lo tanto, el universo no puede originalmente alcanzar su orden y disposición sin algo similar al arte humano (parte II).

Además, observa Philo, si situamos a la mente como la causa primera, con su «desconocida e inexplicable economía», lo que se hace es posponer el

problema:

Estamos obligados a ascender aún más, con el fin de encontrar la causa de esta causa, a la cual has caracterizado como satisfactoria y concluyente… sin embargo ¿cómo podemos satisfacemos a nosotros mismos con respecto a la causa de ese Ser, a quien se le supone que es el Autor de la naturaleza, o, de acuerdo con tu sistema antropomórfico, el mundo ideal, en el cual buscas el mundo material? ¿No tenemos la misma razón para seguir las huellas de este mundo ideal en otro mundo ideal o en un nuevo principio inteligente? Pero si nos detenemos y no seguimos más lejos; ¿por qué ir tan lejos? ¿Por qué no detenernos en el mundo material? ¿Cómo podemos satisfacernos a nosotros mismos sin proseguir hasta el infinito? Y, después de todo, ¿qué satisfacción encontramos en esta infinita progresión? (parte IV).

Cleantes no tiene respuestas satisfactorias para estas preguntas retóricas y lo que sigue es peor. Cleantes insiste en que la mente de Dios es *como la humana* y está de acuerdo cuando Philo añade «la mejor de la mejor». Pero entonces, Philo presiona a Cleantes: ¿es la mente de Dios perfecta, «libre de todo error, equivocación o incoherencia en sus acciones» (parte V)? Hay una hipótesis opuesta que debe ser descartada:

¿Qué sorpresa nos invadiría si lo que encontramos es un estúpido mecánico que ¡mita a los otros y copia un arte, el cual, a través de una larga sucesión de edades, después de múltiples ensayos, errores, correcciones, deliberaciones y controversias, ha sido gradualmente mejorado? Muchos mundos pueden haber sido construidos chapuceramente, a través de una eternidad antes de que su sistema fuera destruido: mucho trabajo perdido, muchos ensayos infructuosos realizados, una lenta pero continuada mejoría conseguida durante infinitas edades de construcción del mundo (parte V).

Cuando Philo presenta esta fantástica alternativa, con su asombrosa anticipación a la perspicacia de Darwin, no la toma en serio excepto como arma para combatir el punto de vista de Cleantes, que defiende la existencia de un Artífice omnisciente. Hume utiliza esta alternativa solamente para señalar lo que él considera limitaciones de nuestro conocimiento: «En tales casos, ¿quién puede determinar dónde está la verdad?; aún más, ¿quién puede conjeturar dónde está la probabilidad?; en medio, un gran número de hipótesis que pueden ser propuestas y un número aún mayor de hipótesis que pueden ser imaginadas» (parte V).

La imaginación corre desbordada y, explotando esta fecundidad, Philo va acorralando a Cleantes, diseñando extrañas y cómicas variaciones de sus propias hipótesis, desafiándolo a que demuestre por qué es su propia versión la que debe ser la preferida. «¿Por qué no pueden ser varias Deidades las que, actuando combinadamente, hayan ideado y estructurado un mundo?… ¿Y por qué no llegar a ser un perfecto antropomorfita? ¿Por qué no afirmar que la Deidad o las Deidades son corpóreas, y tienen ojos, nariz, boca, orejas, etc.?»

(parte V). En un determinado momento, Philo anticipa la hipótesis de Gaia: el universo

presenta una gran semejanza con un animal o cuerpo organizado, y parece actuar con un principio semejante de vida y movimiento. En él una continua circulación de materia no produce trastorno… El mundo, por lo tanto, según infiero, es un animal y la Deidad es el alma del mundo actuando en él y por él (parte VI).

¿O quizá no es el mundo realmente más parecido a un vegetal que a un animal?

De la misma manera que un árbol esparce sus semillas en los campos vecinos y da nacimiento a otros árboles; así el gran vegetal, el mundo, o su sistema planetario, produce dentro de sí ciertas semillas, las cuales diseminadas en el caos que le rodea vegetan en nuevos mundos. Un cometa, por ejemplo, es la semilla de un mundo… (parte VII).

Una posibilidad, en cierto modo, más extravagante:

Los brahmanes aseguran que el mundo nació de una araña infinita, que hiló la totalidad de esta complicada masa en sus intestinos y aniquiló después la totalidad o parte de esta masa, absorbiéndola de nuevo y disolviéndola en su propia esencia. Aquí nos encontramos con una especie de cosmogonía, que nos parece ridícula; ya que una araña es un pequeño y despreciable animal, cuyas operaciones no tomaremos nunca como modelo de todo el universo. Pero aun aquí se encuentra un nuevo tipo de analogía, hasta en nuestro globo terráqueo. Y si hubiera un planeta habitado totalmente por arañas (lo cual es muy posible) esta inferencia aparecería tan natural e irrefutable como aquella que en nuestro planeta adscribe el origen de todas las cosas al diseño y a la inteligencia, como ha explicado Cleantes. Será difícil dar una razón satisfactoria de por qué un sistema ordenado no puede ser hilado en el vientre del mismo modo que lo es en el cerebro (parte VII).

Cleantes resiste valerosamente estos furiosos asaltos, pero Philo descubre grietas fatales en cada versión del argumento que aquél consigue diseñar. Sin embargo, Philo nos sorprende, al final de los *Diálogos*, poniéndose de acuerdo con Cleantes.

La conclusión legítima es que… si no nos satisface llamar *Dios* o *Deidad* a la primera y suprema causa y deseamos cambiar su nombre, ¿qué tal llamarle *Mente* o *Pensamiento*, con la cual es justo suponer que tiene una considerable semejanza? (parte XII).

Philo es seguramente la voz de Hume en los *Diálogos*. ¿Por qué cedió Hume? ¿Acaso tuvo miedo a represalias por parte del sistema? No. Hume se dio cuenta de que él había demostrado que el argumento del diseño era un puente irreparablemente agrietado entre la ciencia y la religión y arregló las cosas para que sus *Diálogos* fueran publicados después de su muerte en 1776, precisamente para salvarse de la persecución. Hume cedió porque justamente *no podía imaginar* otra explicación sobre el origen del manifiesto diseño en la

naturaleza. Hume no pudo comprender cómo la «curiosa adaptación de los medios a los fines», observada en toda la naturaleza, «podía ser debida al azar»; y si no es debida al azar, ¿a qué es debida?

¿Quién, si no un Dios inteligente, podía ser responsable de este diseño de tan alta calidad? Philo es uno de los contrincantes más ingeniosos y con más recursos en cualquier debate filosófico, real o imaginario, y logra algunas admirables brechas en la oscuridad, luchando por una alternativa. En la parte VIII sueña algunas especulaciones que están muy cerca de adelantarse provocadoramente a Darwin (y a algunas más recientes elaboraciones de la teoría de Darwin) en casi un siglo.

En lugar de una presunta materia infinita, como suponía Epicuro, supongamos que es finita. Un número finito de partículas es sólo susceptible de transposiciones finitas. Y esto debe suceder, en una duración eterna, que todo posible orden o posición debe ser logrado un infinito número de veces… Hay un sistema, un orden, una economía de las cosas por los cuales la materia puede conservar una agitación perpetua, que parece esencial y además mantener una constancia en las formas. ¿Quién lo produce? Existe, ciertamente, tal economía: éste es realmente el caso del presente mundo. Por lo tanto, el continuo movimiento de la materia, en menos que infinitas transposiciones, debe producir esta economía u orden; y por su propia naturaleza, este orden, una vez establecido, mantenerse a sí mismo, durante mucho tiempo, si no hasta la eternidad. Pero dondequiera que la materia se encuentra así estabilizada, estructurada y ajustada a un movimiento continuo perpetuo, y todavía conserva una constancia en las formas, su situación debe, por necesidad, tener toda ella la misma apariencia de arte y de invención que observamos al presente… Un defecto en cualquiera de estas partículas destruye la forma; y la materia, de la cual está compuesta, es de nuevo diseminada, impulsada en movimientos irregulares y fermentaciones, hasta que se reúne en alguna otra forma regular…

Supongamos… que la materia fuera lanzada a cualquier posición por una fuerza ciega y sin guía; es evidente que su primera posición debe ser, con roda probabilidad, la más confusa y más desordenada que pueda ser imaginada, sin ninguna semejanza con aquellas obras de la invención humana, en las cuales, además de una simetría de las partes, se descubre un ajustamiento de los medios a los fines y una tendencia a la autopreservación… Supongamos que la fuerza actuante, cualquiera que ella sea, aún continúa en funcionamiento… Así el universo marcharía, a través de muchas edades, en una continua sucesión de caos y desorden. Pero ¿no es posible que esa situación termine por estabilizarse?… ¿No deberíamos esperar esa situación, o incluso estar seguros de ella, a partir de las eternas revoluciones de la materia sin guía, y no podría esto explicar toda la aparente sabiduría e inventiva que existe en el universo?

¡Bien!, parece que algo parecido puede funcionar… pero Hume no puede tomar completamente en serio la atrevida incursión de Philo. Su veredicto final: «Una suspensión total del juicio parece en este momento nuestro único recurso razonable» (parte VIII). Unos cuantos años antes, Diderot había escrito algunas especulaciones que sorprendentemente presagiaban a Darwin: «Yo puedo mantener ante ti… que monstruos aniquilan a otros monstruos en sucesión; que todas las combinaciones defectuosas de materia han desaparecido y que tan sólo han sobrevivido aquellas en las que la organización no incluye alguna importante contradicción, y pueden subsistir

por sí mismas y perpetuarse por sí mismas» (*Carta sobre los ciegos para uso de los que ven*, 1749). Ideas atractivas acerca de la evolución han estado flotando durante milenios aunque, como la mayoría de las ideas filosóficas, mientras que parecen ofrecer una solución a los problemas planteados, no pueden prometer ir más lejos, para abrir nuevas investigaciones o generar sorprendentes predicciones que puedan ser verificadas, o explicar hechos para cuya explicación fueron expresamente concebidos. El carácter revolucionario de la evolución tuvo que esperar hasta que Charles Darwin supo cómo convertir una hipótesis evolucionista en un tejido de explicaciones compuesto por literalmente miles de hechos, a menudo sorprendentes, acerca de la naturaleza conseguido con enorme esfuerzo. Darwin ni inventó la admirable idea de todo este tejido de explicaciones por sí mismo ni la comprendió en su totalidad ni siquiera cuando la formuló. Pero él hizo el monumental trabajo de clarificar la idea, anclándola para que no volviera a flotar, por lo que merece el crédito más que ningún otro. El siguiente capítulo revisa este logro básico de Darwin.

CAPÍTULO 2
Ha nacido una idea

I. *¿Qué hay de especial en las especies?*

Charles Darwin no empezó por preparar la mezcla de un antídoto contra la parálisis conceptual de Locke o aclarar la gran alternativa cosmológica que había simplemente eludido Hume. Una vez que se le ocurrió su gran idea, se dio cuenta de que tendría realmente verdaderas consecuencias revolucionarias, aunque al principio él no trataba de explicar el significado de la vida, ni siquiera sus orígenes. Su objetivo era algo más modesto: deseaba explicar el origen de las *especies*.

En la época de Darwin, los naturalistas habían acumulado montañas de hechos sorprendentes acerca de los seres vivos y habían conseguido sistematizar estos hechos desde varias dimensiones. Dos grandes fuentes de admiración emergían de este trabajo (Mayr 1982). En primer lugar, estaban todos los descubrimientos acerca de las *adaptaciones* de los organismos que habían subyugado al Cleantes de Hume: «Todas estas diversas máquinas, y hasta en sus partes más diminutas, están ajustadas unas a otras con una precisión que provoca la admiración de todos los hombres que alguna vez las han contemplado» (parte II). En segundo lugar, estaba la prolífica *diversidad* de los seres vivos, literalmente millones de diferentes clases de plantas y animales. ¿Por qué había tantos?

La diversidad en el diseño de los organismos era tan llamativa, en algunos aspectos, como la excelencia de estos diseños y aún más llamativos eran los modelos discerníbles dentro de la diversidad. Podían observarse miles de gradaciones y variaciones entre los organismos, aunque había también grandes hiatos entre ellos. Había pájaros y mamíferos que nadaban como peces, pero ninguno con branquias; había perros de muchos tamaños y formas, pero no había perros-gato o perros-vaca o perros con plumas. Los modelos exigían ser clasificados y en los tiempos de Darwin el trabajo de los grandes taxonomistas (que comenzaron por adoptar o corregir la antigua clasificación de Aristóteles) había construido una detallada jerarquía de dos

reinos (plantas y animales) divididos en *phyla*, las cuales se dividían en clases, éstas en órdenes, éstos en familias, éstas en géneros (el plural de «genus») y éstos en especies. Las especies podían ser subdivididas, naturalmente, en subespecies o variedades: los perros *cocker spaniel* y *basset* son diferentes variedades de una única especie de perro, el *Canis familiaris*.

¿Cuántos tipos diferentes de organismos existen? Dado que dos organismos no son exactamente iguales —ni incluso los gemelos univitelinos — hay tantos tipos de organismos como organismos existen, pero parece obvio que las diferencias son de distinto grado, de menor a mayor, o *accidentales* y *esenciales*. Así lo había enseñado Aristóteles, y esta enseñanza era una pizca de filosofía que había penetrado en el pensamiento de todo el mundo, desde cardenales a boticarios y a vendedores ambulantes. Todas las cosas —no exactamente seres vivos— tienen dos clases de propiedades: propiedades esenciales sin las cuales no serían el tipo particular de cosa que son y propiedades accidentales que pueden variar libremente dentro cada tipo. Una pieza de oro puede cambiar su forma *extemporáneamente* y seguirá siendo oro; lo que hace que sea oro es su propiedad esencial, no sus accidentes. A cada tipo le corresponde una esencia. Las esencias eran definitivas y como tales son intemporales, inmodificables y se rigen por la ley del todo-o-nada. Una cosa no puede ser *más bien* plata o *casi* oro, o bien un *semi*-mamífero.

Aristóteles había desarrollado su teoría de las esencias como un mejoramiento de la teoría platónica de las ideas, según la cual cada cosa terrena es como una copia imperfecta o una reflexión de una idea ejemplar o forma que existe eternamente en el reino platónico de las ideas, regidas por Dios. Este reino platónico de las abstracciones no es visible, naturalmente, pero sí era accesible a la mente por medio del pensamiento deductivo. Lo que los geómetras pensaban y los teoremas demostraban, por ejemplo, eran las formas del círculo y del triángulo. Dado que existían las formas del águila y del elefante, una ciencia deductiva de la naturaleza era también merecedora de ser considerada. Pero así como ningún círculo terrestre, no importa lo cuidadosamente que se trace con un compás, o mediante una rueda de alfarero, puede ser uno de los círculos perfectos de la geometría de Euclides, del mismo modo ningún águila puede manifestar perfectamente la esencia de ser águila aunque toda águila se esfuerza por serlo. Todas las cosas existentes tienen una especificación divina, que captura su esencia. La taxonomía de los seres vivos heredada por Darwin era en sí misma descendiente directo, a través de Aristóteles, del esencialismo platónico. En efecto, la palabra

«especie» fue, en cierto modo, una traducción estándar de la palabra griega *eidos* que Platón utilizó para designar la Forma o Idea.

Nosotros, postdarwinistas, estamos tan acostumbrados a pensar en términos históricos acerca del desarrollo de las formas de vida que nos cuesta trabajo recordar que en los días de Darwin las especies de los organismos eran consideradas tan eternas como los triángulos y las circunferencias perfectos de la geometría de Euclides. Sus miembros individuales aparecen y desaparecen, pero las especies permanecen sin alterar e inalterables. Esto formaba parte de una herencia filosófica, pero no era un dogma inútil o deficientemente motivado. Los triunfos de la ciencia moderna, desde Copérnico y Kepler, Descartes y Newton, habían supuesto la aplicación de unas matemáticas precisas al mundo material y esto, aparentemente, requiere la abstracción de las descuidadas propiedades accidentales de las cosas para encontrar sus secretas esencias matemáticas. No importa el color o la forma de una cosa cuando se trata de obedecer la ley de la atracción gravitacional de Newton. Lo único que importa es su masa. De modo similar, la alquimia ha sido sustituida por la química, una vez que los químicos establecieron su credo fundamental: existe un número finito de elementos básicos, *inmutables*, tales como el carbono, el oxígeno, el hidrógeno y el hierro. Estos elementos pueden mezclarse y unirse en combinaciones sin fin a lo largo del tiempo, pero los bloques fundamentales de estas construcciones son identificables por sus propiedades esenciales e inmutables.

La doctrina de las esencias aparece como un poderoso organizador de los fenómenos del mundo en muchas áreas, pero ¿qué es lo verdadero en todos los esquemas clasificatorios que se pueden diseñar? ¿Cuáles son las diferencias esenciales entre colinas y montañas, nieve y aguanieve, mansiones y palacios, violines y violas? John Locke y otros desarrollaron elaboradas doctrinas que distinguían las esencias reales de las esencias *nominales*; las últimas serían simples parásitos de los *nombres* o palabras que elegimos utilizar. Se puede establecer la clasificación que se desee; por ejemplo, un club dedicado a la cría de perros puede votar una lista definitiva de las condiciones necesarias para que un perro sea considerado un genuino *Ourkind Spaniel*, pero ésta sería una esencia nominal, no una esencia real. Las esencias reales pueden descubrirse mediante la investigación científica en la naturaleza interna de las cosas, donde la esencia y el accidente pueden distinguirse conforme a unos principios. Era difícil expresar con precisión lo que eran los principios *principales*, pero con los avances de la química y de la física

marcando el camino a seguir, parece razonable que se puedan definir también marcas de las esencias reales en los seres vivos.

Desde la perspectiva de esta deliciosamente precisa y sistemática visión de la jerarquía de los seres vivos, existe un considerable número de hechos embarazosos y enigmáticos. Estas aparentes excepciones fueron casi tan inquietantes para los naturalistas como el descubrimiento de un triángulo cuyos ángulos no sumaran 180 grados lo hubiera sido para un geómetra. Aunque muchos de los límites taxonómicos están bien marcados y aparentemente sin excepciones, es difícil clasificar las criaturas intermedias, que parecen tener partes que corresponden a más de una esencia. Existen también los curiosos modelos de los órdenes más elevados de hechos compartidos y no compartidos: ¿por qué es el espinazo y no las plumas lo que los pájaros y los peces comparten y por qué no puede ser *criatura con ojos* o el ser *carnívoro* tan importante para una clasificación como *criatura de sangre caliente*? Mientras que los amplios diseños y la mayor parte de las reglas específicas de la taxonomía eran indiscutibles (y por supuesto lo siguen siendo en la actualidad), hubo duras controversias sobre los casos problemáticos. ¿Eran todos los lagartos miembros de la misma especie o de varias especies diferentes? ¿Qué principio de clasificación debe «contar»? Según la famosa imagen de Platón, ¿qué sistema «moldeó la naturaleza en sus articulaciones»?

Antes de Darwin, estas controversias estaban básicamente mal planteadas y no podían propiciar una respuesta sólida y razonada, debido a que no existía el fundamento teórico de *por qué* una clasificación debía ser tenida en cuenta para abordar las relaciones existentes, es decir, las pautas por las que las cosas habían llegado a ser como eran. Hoy las librerías se enfrentan al mismo tipo de problema mal planteado: ¿cómo deben organizarse entrecruzadamente las siguientes categorías: *best-sellers*, ciencia-ficción, terror, jardinería, biografía, novelas, colecciones, deportes, libros lujosos? Si el terror es un género de ficción, entonces las historias de terror verdaderas presentan un problema. ¿Deben ser consideradas como ficción todas las novelas? En este caso el librero no puede respetar la propia definición de Truman Capote y considerar *A sangre fría* (1965) como una novela de no ficción, aunque el libro no puede ubicarse confortablemente entre las biografías o los libros de historia. ¿En qué sección de la librería debe colocarse el libro que está leyendo? Obviamente no hay un método correcto para clasificar libros; en este terreno sólo encontraremos esencias nominales. Sin embargo, muchos naturalistas estaban convencidos, por principios generales, que había esencias reales entre las

categorías de su sistema natural de seres vivos. Como Darwin puntualizó, «ellos creen que esto revela el plan del Creador; pero a menos que se especifique si es orden en el tiempo y en el espacio, o qué otra cosa se quiere significar con el plan del Creador, me parece que nada se añade a nuestro conocimiento» (*El origen de las especies*).

A veces los problemas de la ciencia se hacen más fáciles si se les añaden complicaciones. El desarrollo de la ciencia de la geología y el descubrimiento de fósiles de especies manifiestamente extinguidas aumentó la curiosidad de los taxonomistas hasta confundirlos, pero estas curiosidades fueron también las piezas del rompecabezas que permitió a Darwin, trabajando paralelamente con centenares de científicos, descubrir la clave de su solución: las *especies* no eran eternas ni inmutables; habían evolucionado a lo largo del tiempo. A diferencia de los átomos de carbono, que, hasta donde sabemos, se han mantenido siempre exactamente en la misma forma que ahora exhiben, las especies han nacido en el tiempo, pueden cambiar a lo largo del tiempo y, a su vez, pueden dar origen a nuevas especies. La idea, en sí misma, no era nueva; diferentes versiones habían sido ya discutidas, remontándonos a los antiguos griegos. Pero existía una poderosa predisposición platónica contra ella: las esencias eran inalterables y una cosa no puede cambiar su esencia, del mismo modo que no podían nacer nuevas esencias, excepto, naturalmente, por deseo de Dios dentro de episodios de una creación especial. Los reptiles no podían *convertirse* en pájaros, del mismo modo que el cobre no podía transformarse en oro.

No es fácil hoy simpatizar con esta convicción, pero el esfuerzo para lograrlo puede verse fomentado por la fantasía: considere el lector cuál sería nuestra actitud hacia una teoría que se propusiera demostrar que el número 7 fue, hace mucho, mucho tiempo, un número par y que gradualmente había adquirido su carácter impar a través de ciertas disposiciones mediante las cuales había intercambiado algunas propiedades con los ancestros del número 10 (el cual fue antaño un número primo). Una soberana tontería, naturalmente. Darwin sabía que una actitud paralela estaba profundamente arraigada entre sus contemporáneos y que tendría que esforzarse mucho para vencerla. Sin duda, Darwin admitía, más o menos, que las autoridades más veteranas de su tiempo tenderían a mantenerse tan inmutables como las especies en las que ellos creían, así que en la conclusión de su libro solicitaba el apoyo de los jóvenes lectores: «Cualquiera que llegue a creer que las especies son mutables prestará un buen servicio expresando conscientemente

su convicción; sólo de esta manera será posible levantar la losa de prejuicios que aplastaba a este tema» (*El origen de las especies*).

Aún hoy, la demolición del esencialismo por parte de Darwin no ha sido completamente asimilada. Por ejemplo, en estos días se discute mucho en filosofía acerca de los «tipos naturales», un antiguo término del filósofo W. V. O Quine [1969] muy cautelosamente resucitado para uso limitado al distinguir las buenas categorías científicas de las malas. Pero en los escritos de otros filósofos, la expresión «tipo natural» es a menudo un modo de vestir de oveja al lobo de la esencia real. La presión esencialista se ejerce todavía entre nosotros y no siempre por malas razones. La ciencia aspira a «moldear la naturaleza en sus articulaciones» y a menudo parece como si necesitáramos a las esencias, o algo parecido a las esencias, para hacer este trabajo. Sobre este punto, los dos grandes reinos del pensamiento filosófico, el platónico y el aristotélico, están de acuerdo. Pero la mutación darwiniana, que al principio parecía una nueva vía para pensar acerca de las clases en biología, puede extenderse, como hemos visto, a otros fenómenos y a otras disciplinas. Existen continuos problemas tanto dentro como fuera de la biología que rápidamente se solucionan una vez que hemos adoptado la perspectiva darwiniana sobre lo que hace que una cosa sea lo que es, pero persiste la resistencia a esta idea ligada a la tradición.

2. *Selección natural. Algo tremendamente difícil de creer*

> Cuesta muchísimo creer que la cola de un pavo real fue formada de esta manera, pero, una vez creído esto, yo creo en el mismo principio, algo modificado, aplicado al hombre.
>
> *Charles Darwin*
> (Carta citada en Desmond y Moore, *Darwin*)

El proyecto de Darwin en *El origen de las especies* puede ser dividido en dos partes: probar que las especies modernas eran descendientes modificados de especies más tempranas —especies que habían evolucionado— y demostrar *cómo* había ocurrido este proceso de «descenso con modificación». Si Darwin no hubiese tenido la visión de un mecanismo, la selección natural, mediante el cual esta casi inconcebible transformación histórica pudo haberse conseguido, probablemente no se hubiera sentido motivado para reunir todas las evidencias circunstanciales como realmente ocurrió. Hoy podemos fácilmente imaginarnos demostrando el primer caso de Darwin —el hecho histórico del

descenso con modificación— con total independencia de cualquier consideración sobre la selección natural o naturalmente de cualquier otro mecanismo para conseguir estos brutales hechos, pero para Darwin la idea del mecanismo era, al mismo tiempo, la licencia de caza que necesitaba y la guía inquebrantable que le condujera a las preguntas correctas que tenía que hacer[6].

La idea de la selección natural no fue, en sí misma, una milagrosa nueva creación de Darwin, sino, más bien, el fruto de ideas anteriores que habían sido duramente discutidas durante años y generaciones (para un excelente relato de esta historia intelectual, puede leerse a R. Richards 1987). Entre estas ideas progenitoras, la principal fue una reflexión de Darwin sobre el *Ensayo sobre el principio de la población* (1798) de Thomas Malthus, quien argumentaba que la explosión de la población y el hambre eran inevitables, dada la excesiva fertilidad de los seres humanos, a menos que se tomaran medidas drásticas. La desconsoladora visión malthusiana de las fuerzas políticas y sociales que podrían actuar para controlar la superpoblación humana podía haber aderezado el pensamiento de Darwin (e indudablemente ha aderezado los triviales ataques políticos de muchos antidarwinianos), pero la idea que Darwin necesitaba de Malthus era puramente lógica. No tiene nada que ver con la ideología política, y puede ser expresada en términos muy abstractos y generales.

Supongamos un mundo en el cual los organismos tengan muchos descendientes. Dado que los descendientes tendrán, a su vez, muchos descendientes, la población crecerá y crecerá («geométricamente») hasta que inevitablemente, tarde o temprano —sorprendentemente temprano, de hecho —, se habrá incrementado demasiado para los recursos disponibles (de alimentos, de espacio, de cualquier cosa que los organismos necesiten para sobrevivir lo suficiente para reproducirse). En este punto, cuando esto suceda, no todos los organismos tendrán descendencia. Muchos morirán sin ella. Esto era lo que Malthus señalaba como la inevitabilidad matemática de tal quiebra en cualquier población de reproductores a largo plazo —*seres* humanos, animales, plantas (o cualquier idea de ciencia ficción, aunque estas fantásticas posibilidades no fueran discutidas por Malthus). Aquellas poblaciones que se reproducen por debajo de la tasa de reemplazamiento están abocadas a la extinción a menos que inviertan su tendencia. Las poblaciones estables durante largos períodos de tiempo, lo logran manteniendo una tasa de sobreproducción de descendientes que compensa las pérdidas ocasionadas por una alta mortalidad. Esto es obvio, quizá, para las moscas y otros prodigiosos

reproductores, pero Darwin remachó el clavo con un cálculo de su propia cosecha: «El elefante se considera el reproductor más lento de todos los animales conocidos y me he tomado algunas molestias para calcular su probable tasa mínima de crecimiento natural… al final de siglo V habría vivos quince millones de elefantes, descendientes del primer par» (*El origen de las especies*)[7] Dado que los elefantes existen desde hace millones de años, podemos asegurar que solamente una fracción de los elefantes nacidos en cualquier período tiene sus propios descendientes.

De este modo, el estado normal para cualquier tipo de reproductor es aquel en el que el número de descendientes producidos en cualquier generación es mayor que el de los que a su vez se reproducirán en la siguiente. En otras palabras, casi siempre es tiempo de quiebra en la progresión de la población[8]. Cuando se produce esta quiebra, ¿cuáles de los probables padres «ganarán»? ¿Será una lotería sin trampas en la que todos los organismos tendrán la misma oportunidad de encontrarse entre los pocos que se van a reproducir? En un contexto político, es en este punto donde se introducen temas odiosos acerca de poder, privilegio, injusticia, traición, lucha de clases y similares, aunque podemos llevar esta observación más allá de su lugar de nacimiento político y considerar en abstracto, como hizo Darwin, lo que sucedería o debería suceder en la naturaleza. Darwin añadió dos aspectos lógicos al descubrimiento de Malthus: el primero era que cuando se produce la quiebra en el crecimiento de la población, si había variaciones significativas entre los contendientes, entonces cualquier ventaja mostrada por alguno de ellos predispondría inevitablemente la muestra que se reproducía. Por leve que fuese la ventaja en cuestión, si realmente era una ventaja (de modo que no pasara inadvertida a la naturaleza), inclinaría la balanza del lado del que la tuviera. El segundo era que si había un «fuerte principio de herencia» —si los descendientes tendían a parecerse más a sus padres que a los contemporáneos de sus padres— las predisposiciones creadas por las ventajas, aunque pequeñas, llegarían a amplificarse con el tiempo, creando tendencias que crecerían indefinidamente. «Han nacido más individuos de los que pueden sobrevivir. Un gramo en la balanza determinará qué individuo vivirá y qué individuo morirá, qué variedad de especie incrementará su número y qué especie decrecerá o finalmente se extinguirá» (*ibidem*).

Lo que Darwin observó fue que si se suponen estas escasas condiciones generales aplicables a las épocas de crisis poblacional —condiciones para las que aportó pruebas suficientes— el proceso resultante conduciría *necesariamente* en la dirección de que individuos de la futuras generaciones

tenderían a estar mejor equipados para tratar con los problemas de limitación de recursos de lo que habían estado los individuos de la generación de sus padres. Esta idea fundamental —la peligrosa idea de Darwin, la idea que genera tantas perspectivas, alboroto, confusión y ansiedad— es, pues, realmente muy simple. Darwin la había resumido en dos largas sentencias al final del capítulo 4 de *El origen de las especies*:

> Si durante el largo devenir de los tiempos y bajo condiciones variables de vida, los seres orgánicos varían en las diversas partes de su organización, yo pienso que esto no puede ser discutido; si ha tenido lugar, debido a la elevada capacidad de incremento geométrico de cada especie, en alguna edad, estación o año, una intensa lucha por la vida, y esto ciertamente no puede ser disputado; entonces, considerando la infinita complejidad de las relaciones de todos los seres orgánicos entre sí y de sus condiciones de existencia, causa de una infinita diversidad en la estructura, constitución y hábitos, que es ventajosa para ellos, yo creo que sería un hecho muy extraordinario que no ocurriera variación alguna en el propio bienestar de cada ser, del mismo modo como han ocurrido variaciones útiles para el hombre. Pero si ocurren variaciones útiles para todos los seres orgánicos, seguramente los individuos así caracterizados tendrán las mejores posibilidades de ser preservados en la lucha por la vida; y apoyándose en el fuerte principio de la herencia, tenderán a producir descendientes caracterizados de modo similar. El principio de la preservación ha sido llamado, en aras de la brevedad, selección natural.

Ésta ha sido la gran idea de Darwin, no la idea de la evolución sino la idea de la evolución por selección natural, una idea que él mismo nunca formuló con suficiente rigor y detalle para probarla, aunque presentó un brillante informe. Las dos próximas secciones del libro se concentrarán en los hechos curiosos y cruciales de esta sumaria declaración de Darwin.

3. *¿Explicó Darwin el origen de las especies?*

> Darwin luchó con brillantez y consiguió triunfar frente al problema de la adaptación, pero obtuvo un éxito limitado con el problema de la diversidad, aunque el título de su libro hace referencia a éste su fallo relativo: el origen de las especies.
>
> Stephen Jay Gould,
> «*The Confusion over the Evolution*»

> De este modo, el gran hecho en la historia natural que es el de la subordinación de un grupo bajo otro grupo, el cual, dada su familiaridad, no siempre nos ha llamado la atención suficientemente, queda, a mi juicio, completamente explicado.
>
> Charles Darwin, *El origen de las especies*

Conviene subrayar que el resumen de Darwin no hace mención del proceso de formación de las especies. Trata de la adaptación de los organismos, de la *excelencia* de su diseño, no de la diversidad. Por otra parte, frente a esto, en

su resumen Darwin considera la diversidad de las especies *como algo asumido*: «la infinita complejidad de las relaciones de todos los seres orgánicos entre sí con sus condiciones de existencia». Lo que determina esta prodigiosa (si no realmente infinita) complejidad es la presencia al mismo tiempo (y compitiendo por el mismo espacio vital) de tantas y tan diversas formas de vida, con necesidades y estrategias muy diferentes. Darwin ni siquiera pretendía ofrecer una explicación del origen de las *primeras* especies, o de la propia vida; Darwin comienza en el medio, dando por supuesta la existencia de muchas y muy diferentes especies con muy diversos talentos ya presentes y subraya que, al comenzar desde este punto medio, el proceso que ha descrito inevitablemente asentaría y diversificaría los talentos de las especies ya existentes. Pero ¿seguiría este proceso creando especies? El resumen de Darwin guarda silencio sobre este punto, pero su libro no. En realidad, Darwin imaginó su idea dando explicaciones de un solo golpe y, al mismo tiempo, planteaba grandes interrogaciones. La generación de las adaptaciones y la generación de la diversidad eran diferentes aspectos de un único y complejo fenómeno, y la visión unificadora, para Darwin, era el principio de la selección natural.

La selección natural produciría inevitablemente la *adaptación*, como el resumen de Darwin deja claro, y bajo las circunstancias correctas, como él argumentaba, la acumulación de adaptaciones crearía las especies. Darwin sabía bien que la explicación de las variaciones no era suficiente para explicar la formación de las especies. Los criadores de animales, a los que Darwin extraía con tesón sus conocimientos, sabían cómo obtener la *variedad* dentro de una especie, pero aparentemente nunca habían creado una nueva *especie* y se burlaban de la idea de que las diferentes especies que ellos criaban podrían haber tenido un ancestro común. «Preguntad, como yo lo he hecho a un famoso criador de ganado vacuno Hereford, si este ganado podría ser descendiente del conocido como *longhorn* y se reirá desdeñosamente». ¿Por qué? Porque «aunque ellos saben bien que en cada raza las variaciones son leves, y ganan sus premios seleccionando estas leves diferencias, todavía ignoran todos los principios generales y rehúsan sintetizar en sus mentes las ligeras diferencias acumuladas durante muchas sucesivas generaciones» (*El origen de las especies*).

La posterior diversificación en especies ocurriría, según Darwin, debido a que si llegara a haber una variedad de habilidades heredables o de aprovisionamientos en una población (de una única especie), estas diferentes habilidades o equipamientos tenderían a generar diferentes «utilidades» en los

distintos subgrupos de población, y estas subpoblaciones, a su vez, tenderían a diverger evolutivamente, prosiguiendo cada una su camino hacia la excelencia, hasta que eventualmente se produjera una completa separación en dos caminos distintos. ¿Por qué, se preguntaba Darwin, conduciría esta divergencia a la separación o acumulación de las variaciones, en lugar de mantenerse, más o menos, una continuada distribución de las leves diferencias abierta a toda la población? El mero aislamiento geográfico era una parte de la respuesta; cuando una población quedaba separada por accidentes geológicos o climáticos masivos, o por la fortuita emigración a un espacio aislado, como una isla, esta discontinuidad con el entorno tendría eventualmente su correspondencia en una discontinuidad en las útiles variaciones observadas en las dos poblaciones. Una vez que la discontinuidad se fortalecía, se iría autorreforzando dicha vía de separación hasta concluir en el establecimiento de dos especies distintas. Otra idea de este proceso, aunque diferente, se fundamentaba en la lucha dentro de cada especie (lucha intraespecífica), lucha en la que tendería a funcionar el principio de que «el ganador se lo lleva todo»:

Debe recordarse que la competición será generalmente más dura entre aquellas formas que están relacionadas más próximamente con otras en hábitos, constitución y estructura. De aquí que todas las formas intermediarias entre los estadios precoces y los más tardíos de una especie, es decir, entre los estadios que han sido menos mejorados y los más mejorados, del mismo modo que los progenitores de la especie original, tienden en general, a extinguirse (*ibidem*).

Darwin formuló otras especulaciones ingeniosas y plausibles sobre cómo y por qué la inexorable acción de la selección natural crearía realmente límites entre las especies, aunque siguen siendo especulaciones. Ha costado un siglo de trabajo sustituir las brillantes pero inconclusas meditaciones de Darwin sobre los mecanismos de formación de las especies con explicaciones que son, en cierto grado, demostrables. Persisten aún controversias sobre los mecanismos y los principios que rigen la formación de las especies, así que, en cierto sentido, ni Darwin ni los darwinianos que le han secundado han explicado el origen de las especies. Como ha señalado el genetista Steve Jones [1993], si Darwin quisiera hoy publicar su obra maestra bajo el conocido título «se habría encontrado con dificultades con el Registro de Patentes y Marcas, debido a que si hay algo de lo que no se ocupa *El origen de las especies* es precisamente del origen de las especies. Darwin no sabía nada de genética. Ahora sabemos bastantes cosas, y aunque la vía por la que se inician las especies es un misterio, poseemos ya bastantes detalles».

Pero el hecho de la formación de las especies («especiación») es incontestable, como Darwin demostró elaborando un irrefutable alegato basado en cientos de hechos cuidadosamente estudiados y argumentados. Así es como se originan las especies: a través del «descenso con modificaciones» desde las especies primarias y no por una creación especial. De modo que, en otro sentido, Darwin explicó indudablemente el origen de las especies. Cualquiera que sea el mecanismo operativo, el proceso comienza claramente con la aparición de variedades dentro de una especie y termina, después de la acumulación de modificaciones, con el nacimiento de una nueva especie en la descendencia evolutiva. Lo que comienza como una variedad «bien definida» se convierte gradualmente en la «dudosa categoría de la subespecie; pero nosotros sólo tenemos que suponer que los pasos en el proceso de modificación son más numerosos o de mayor cuantía, para convertir estas… formas en las bien definidas especies» (*ibidem*).

Conviene subrayar que Darwin es muy cuidadoso al describir el resultado eventual como la creación de especies «bien definidas». Eventualmente, dice Darwin, la divergencia llega a ser tan grande que no hay razón para negar que estamos ante dos especies distintas, no simplemente ante dos variedades diferentes. Pero Darwin no pretende jugar al tradicional juego de definir lo que es la diferencia «esencial»:

> Se observará que yo me refiero con el término especie, como una expresión arbitraria en aras de la conveniencia, a un conjunto de individuos que se asemejan mucho entre ellos y esto no es esencialmente diferente del término variedad, que se aplica a formas menos distintas y más fluctúan tes (*ibidem*).

Una marca estándar —entre otras— de la diferencia de las especies, como Darwin reconoció sin reservas, es el aislamiento reproductivo, de tal modo que no hay cruzamientos reproductivos. Es el cruzamiento el que reúne los grupos separados, mezclando sus genes y «frustrando» el proceso de la formación de especies. No es, naturalmente, que algo desee que se produzca la formación de una especie (Dawkins 1986a:327), pero si el irreversible divorcio que marca la formación de especies tiene que suceder, es necesario que esté precedido por un período de ensayo de esta separación, en el cual el cruzamiento reproductivo se suspende, por la causa que sea, de modo que los grupos separados pueden acentuar la separación. El criterio de aislamiento reproductivo es vago en su definición. ¿Pertenecen los organismos a especies diferentes cuando *no pueden* entrecruzarse o precisamente cuando *no* se entrecruzan? Los lobos, los coyotes y los perros se consideran de diferentes especies y, sin embargo, es posible el cruzamiento entre ellos y sus

descendientes no son, por lo general, estériles, al contrario de las mulas (que descienden de caballo y asno). El perro salchicha y el perro lobo irlandés se consideran de la misma especie, pero a menos que sus propietarios utilicen algún procedimiento artificial se encuentran tan aislados reproductivamente como los murciélagos de los delfines. El ciervo de cola blanca de Maine no se cruza de hecho con el ciervo de cola blanca de Massachusetts, dado que no viajan a tanta distancia, pero si son transportados y puestos en contacto, sí se cruzarían y, naturalmente, se consideran de la misma especie.

Para terminar, un ejemplo biológico que parece encargado especialmente para filósofos: veamos lo que sucede con las gaviotas argénteas que viven en el hemisferio norte, en una extensión cuyos límites forman un círculo alrededor del Polo Norte.

> Cuando observamos la gaviota argéntea desplazándose hacia el oeste, desde Gran Bretaña a Norteamérica, vemos gaviotas que son reconocibles como tales gaviotas argénteas, a pesar de ser ligeramente distintas de la forma británica. Podemos seguir su pista, mientras que su apariencia cambia gradualmente, hasta Siberia. Al llegar a este punto, la gaviota se parece más a la variedad que en Gran Bretaña se denomina pequeña gaviota de manto negro. Desde Siberia, cruzando Rusia, hasta el norte de Europa la gaviota se va pareciendo cada vez más a la pequeña gaviota de manto negro británica. Finalmente, en Europa se cierra el círculo; las dos formas geográficamente extremas se reúnen para formar dos especies totalmente distintas: la gaviota argéntea y la pequeña gaviota de manto negro pueden distinguirse por su aspecto y no se cruzan de manera natural (Mark Ridley 1985:5).

Existen ciertamente especies «bien definidas» —el propósito del libro de Darwin es explicar su origen— pero él nos desilusiona cuando trata de encontrar una definición fundamentada en principios sobre el concepto de especie. Las variedades, insiste Darwin, son «especies incipientes» y lo que normalmente convierte dos variedades en dos especies no es la *presencia* de algo (una nueva esencia para cada grupo, por ejemplo) sino la *ausencia* de algo: los casos intermedios, que existían —que podría decirse son escalones necesarios—, pero que con el tiempo se han extinguido, dejando dos grupos que están de *hecho* aislados desde el punto de vista reproductivo, y que presentan asimismo características diferentes.

El origen de las especies presenta un alegato extraordinariamente persuasivo a favor de la primera tesis de Darwin, es decir, el hecho histórico de la evolución como la causa del origen de las especies; y un provocador informe a favor de la segunda tesis, la que defiende que el mecanismo fundamental responsable de la «descendencia con modificación» era la selección natural**[9]**. Los lectores sensatos del libro simplemente no pueden dudar que las especies han evolucionado a través de los eones que poseían, según Darwin, aunque era difícil de vencer un escrupuloso escepticismo

acerca de la capacidad del mecanismo propuesto para el proceso de la selección natural. Los años transcurridos han incrementado el nivel de confianza en ambas tesis, pero no han borrado la diferencia (Ellegard 19 58 presentó un valioso relato de esta historia). La evidencia de la evolución se extendió no solamente a partir de datos procedentes de la geología, la paleontología, la biogeografía y la anatomía (fuentes principales de Darwin) sino también desde la biología molecular y de todas las ramas de las ciencias de la vida. Para decirlo de manera terminante pero honrada, quien dude hoy de que la variedad de la vida en este planeta fue producida por un proceso de evolución es simplemente un ignorante, un ignorante imperdonable en un mundo en el que tres de cada cuatro personas han aprendido a leer y escribir. Sin embargo, las dudas acerca del poder de la idea de Darwin de la selección natural para explicar este proceso evolutivo todavía son intelectualmente respetables, aunque el peso de la prueba para tal escepticismo ha llegado a ser inmenso, como veremos enseguida.

Así, mientras Darwin dependía de su idea del mecanismo de la selección natural para inspirar y guiar su investigación sobre la evolución, el resultado final invierte el orden de dependencia: Darwin demostró de forma tan convincente que las especies *tenían* que haber evolucionado que pudo entonces darle la vuelta y utilizar este hecho para apoyar su idea más radical, la selección natural. Darwin había descrito un mecanismo o proceso que, de acuerdo con sus argumentos, *podía* haber producido todos estos efectos. Los escépticos se encontraron ante un reto: ¿podían demostrar que los argumentos de Darwin eran erróneos? ¿Podían demostrar que la selección natural era incapaz de producir estos efectos?[10] O bien ¿podrían describir otro proceso que consiguiese estos efectos? ¿Qué *otra* explicación puede darse para la evolución si no es ésta?

El predicamento de Hume fue restablecido por este reto. Hume cedió en su momento porque no pudo imaginar que algo que no fuera un Artífice Inteligente pudiera ser la causa de las adaptaciones que cualquiera podía observar. O dicho más exactamente, el Philo de Hume imaginaba varias alternativas diferentes, pero *Hume no encontró el camino para tomar en serio estas imaginaciones*. Darwin describió cómo un artífice no inteligente podía producir estas adaptaciones en el transcurso de largos períodos de tiempo, y probó que verdaderamente habían ocurrido muchos de los estadios intermediarios necesarios para el proceso propuesto. Ahora el reto a la imaginación se invirtió: dados todos los signos reveladores del proceso histórico que Darwin había descubierto —todas las pinceladas del artista,

podría decirse—, ¿alguien puede imaginar cómo otro proceso que no fuese la selección natural podría haber producido estos efectos? Tan completa ha sido la revocación del peso de la prueba que los científicos a menudo se ven a sí mismos como la imagen especular de la difícil situación en que se encontró Hume. Cuando se enfrentan *prima facie* con una objeción poderosa e irrechazable a la selección natural (nos ocuparemos de los alegatos más fuertes a su debido tiempo) tienden a razonar de este modo: yo no sé (todavía) cómo refutar esta objeción, o vencer esta dificultad, pero dado que no puedo imaginar que otra vía que no sea la selección natural pueda causar estos efectos, tengo que asumir que la objeción es espuria: *de algún modo* la selección natural debe ser suficiente para explicar estos hechos.

Antes de que alguno se lance sobre esta cuestión y me diga que acabo de conceder que el darwinismo es una improbable fe en tanto que religión natural, debe tener en cuenta que hay una diferencia fundamental: habiendo declarado su fidelidad a la selección natural, estos científicos han procedido entonces a cargar con el peso de demostrar cómo superar las dificultades de su tesis y, con el tiempo, han tenido éxito ante este reto. En el proceso, la idea fundamental de Darwin de la selección natural ha sido articulada, extendida, aclarada, cuantificada y profundizada en muchas vías, haciéndose más fuerte cada vez que vencía un reto. Con cada éxito creció en los científicos la convicción de que se encontraban en el camino correcto. Es razonable pensar que una idea que en el fondo fuera falsa seguramente habría sucumbido ante tal serie de ataques. Ésta no es una prueba concluyente, naturalmente, sino una consideración persuasiva. Uno de los objetivos de este libro es explicar por que la idea de la selección natural aparece como un claro ganador, aunque todavía existan controversias no resueltas acerca de cómo se pueden tratar estos fenómenos.

4. *La selección natural como un proceso algorítmico*

> ¿Qué límite puede ponerse a este poder, que ha venido actuando a lo largo de todas edades del mundo, controlando rigurosamente la constitución total, la estructura y los hábitos de cada criatura, favoreciendo lo bueno y rechazando lo malo? Yo no puedo vislumbrar límite alguno a este poder que tan lenta y bellamente va adaptando cada forma a las más complejas relaciones de la vida.
>
> Charles Darwin, *El origen de las especies*

El segundo punto que conviene destacar en el resumen de Darwin es que él presenta su principio como deducible mediante un argumento formal: si se cumplen las condiciones, queda *asegurado* un determinado resultado[11]. Aquí está de nuevo el resumen de Darwin con algunos términos claves en cursiva.

Si durante el largo devenir de los tiempos y bajo las cambiantes condiciones de vida, los seres orgánicos varían en las diversas partes de su organización, yo pienso que esto no puede ser discutido; si ha habido, a causa de la elevada progresión geométrica en el incremento de cada especie, en alguna edad, estación o año, una intensa lucha por la vida, esto ciertamente no puede ser discutido; *entonces*, si se considera la infinita complejidad de las relaciones de todos los seres orgánicos entre sí y de sus condiciones de existencia, que causa una infinita diversidad en la estructura, constitución y hábitos, que es ventajosa para ellos, *yo creo que sería un hecho muy extraordinario que no ocurriera variación alguna en el propio bienestar de cada ser*, del mismo modo como han ocurrido variaciones útiles para el hombre. Pero si ocurren variaciones útiles para todos los seres orgánicos, *seguramente* los individuos así caracterizados tendrán las mejores posibilidades de ser preservados en la lucha por la vida; y apoyándose en el fuerte principio de la herencia, tenderán a producir descendientes caracterizados de modo similar. El principio de la preservación ha sido llamado, en aras de la brevedad, selección natural (*ibidem*).

El argumento deductivo básico es corto y atractivo, aunque el propio Darwin describió el *El origen de las especies* como «un largo argumento», que consiste en dos tipos de demostraciones: la demostración lógica de que un cierto *tipo* de proceso tendría necesariamente que abocar a un cierto tipo de resultado y la demostración empírica de que se han dado en la naturaleza las condiciones requeridas para este tipo de proceso. Darwin refuerza su demostración lógica con experimentos mentales —«casos imaginarios»— que muestran cómo la conjunción de estas condiciones puede realmente dar lugar a los efectos que él trataba de explicar, pero la totalidad de su argumento se extiende a lo largo de todo el libro debido a que presenta una gran riqueza de detalles empíricos conseguidos con mucho esfuerzo para convencer al lector de que estas condiciones se habían dado repetidamente.

Stephen Jay Gould [1985] nos deja entrever la importancia de esta característica del argumento de Darwin en una anécdota acerca de Patrick Matthew, un naturalista escocés que, como hecho histórico curioso, había entrevisto la idea de Darwin sobre la selección natural muchos años antes, en 1831, en un apéndice de su libro *Naval Timber and Arboriculture*. En la estela del ascenso a la fama de Darwin, Matthew publicó una carta (¡en el *Gardener’s Chronicle*¡)[12] proclamando su prioridad, que Darwin graciosamente le concedió, excusando su ignorancia, pero haciendo notar que la revista que había escogido era prácticamente desconocida. Respondiendo a la apología de Darwin, Matthew escribió:

Para mí la concepción de esta ley de la naturaleza llegó intuitivamente como un hecho evidente por sí mismo, casi sin un esfuerzo de concentración mental. El señor Darwin parece tener más mérito en el descubrimiento del que yo he tenido, ya que para mí esto no ha sido un descubrimiento. El señor Darwin parece haberlo conseguido mediante un razonamiento inductivo, lenta y cautelosamente, para continuar su camino sintéticamente, de hecho en hecho; mientras que para mí fue suficiente echar una mirada general al esquema de la naturaleza para estimar que esta selecta producción de especies era un hecho reconocible *a priori*, un axioma, que tan sólo necesita ser expuesto para que sea admitido por mentes sin prejuicios con la suficiente capacidad de captación (citado en Gould 1985:345-346).

Sin embargo, mentes sin prejuicios pueden resistir bastante bien una nueva idea a causa de un sano conservadurismo. Los argumentos deductivos son notoriamente traicioneros; lo que parece «fundado en la razón» puede ser traicionado por un detalle que ha pasado inadvertido. Darwin se dio cuenta de que solamente una cuidadosa y continuada revisión de la evidencia sobre el histórico proceso que él postulaba, persuadiría o podría persuadir a los científicos a abandonar sus convicciones tradicionales y aceptar su revolucionaria visión, incluso si ésta era «deducible de sus primeros principios».

Desde el principio, algunos consideraron la novedosa mezcla de naturalismo detallado y razonamiento abstracto acerca de los procesos de Darwin como un híbrido dudoso e inviable. Esta idea tenía un extraordinario aire de plausibilidad, pero así sucede con muchas de esas recetas para hacerse rico rápidamente que terminan por convertirse en un tremendo fiasco. Comparémoslo con el siguiente principio del mercado de valores: compra barato y vende caro. Esta es la garantía para hacer fortuna. No fracasará en su empeño de hacerse rico *si* sigue este consejo. ¿Por qué no funciona? Sí que funciona, para cualquiera que tenga la suficiente fortuna para actuar de acuerdo con el citado principio, pero ¡ojo!, no hay modo de determinar cuándo se han conjugado las condiciones adecuadas hasta que ya es demasiado tarde para actuar. Darwin estaba ofreciendo a un mundo escéptico lo que podríamos llamar una receta para hacerse rico *lentamente*, un esquema para crear el diseño a partir del caos, sin la ayuda de la mente.

El poder teórico del abstracto esquema de Darwin se debía a varios hechos que él había identificado firmemente y apreciado mejor que muchos de sus partidarios, pero le faltaba la terminología para describirlo explícitamente. En la actualidad podemos englobar estos hechos bajo un sencillo término. Lo que

Darwin había descubierto era el poder de un *algoritmo*. Un algoritmo es un tipo de proceso formal que puede llegar a producir —de forma lógica— un resultado determinado siempre que se le haga funcionar o sea puesto en marcha. Los algoritmos no son nuevos y no eran nuevos en los días de Darwin. Muchos procedimientos aritméticos que nos son familiares, tales como las divisiones largas o el balance del talonario de cheques, son algoritmos y lo son también los procedimientos de tomas de decisiones para jugar perfectamente a tres en raya y para poner una lista de palabras en orden alfabético. Lo relativamente nuevo —si se nos permite una valiosa visión retrospectiva del descubrimiento de Darwin— es la reflexión teórica llevada a cabo por los matemáticos y los lógicos acerca de la naturaleza y el poder de los algoritmos en general, un desarrollo del siglo XX que ha conducido al nacimiento del ordenador, el cual ha conducido a su vez a una más profunda y más viva comprensión de los poderes de los algoritmos en general.

El término *algoritmo* deriva, a través del latín (*algorismus*), del nombre del matemático persa Mûusâ al-Khowârizm, cuyo libro sobre los procedimientos aritméticos, escrito alrededor del año 835 d. C., fue traducido al latín en el siglo XII por Adelardo de Bath o Robert de Chester. La idea de que un algoritmo es un procedimiento seguro y algo «mecánico» se ha mantenido durante siglos, pero han sido los trabajos pioneros de Alan Turing, Kurt Gödel y Alonzo Church en los años treinta los que han contribuido a nuestra comprensión de este término. Tres hechos clave de los algoritmos son importantes para nosotros, y cada uno de ellos es algo difícil de definir. Sin embargo, cada uno de estos hechos ha dado origen a confusiones (y ansiedades) que continúan acosando nuestro pensamiento acerca del descubrimiento revolucionario de Darwin, de modo que tendremos que revisar y reconsiderar estas definiciones introductorias varias veces antes de proseguir:

1) *Sustrato de neutralidad*: el procedimiento de la llamada división larga —en la que se escriben los productos parciales— funciona igualmente bien con lápiz o pluma, papel o pergamino, luces de neón o trazos en el cielo con el humo de un avión, utilizando el sistema de símbolos que se desee. El poder del procedimiento se debe a su estructura lógica, y no a los poderes causales de los materiales utilizados en su realización, justamente el tiempo necesario para que aquellos poderes causales permitan seguir exactamente los pasos prescritos.

2) *Estupidez subyacente*: mientras que el diseño global del procedimiento puede ser brillante, o producir resultados brillantes, cada paso constituyente

así como la transición entre los pasos es absolutamente simple. ¿Cómo de simple? Lo suficientemente simple para que un idiota obediente lo realice o para que un dispositivo mecánico lo lleve a cabo. La analogía estándar que se encuentra en los libros de texto hace notar que los algoritmos son como *recetas*, diseñadas para ser aplicadas por cocineros *novatos*. Un libro de recetas escrito por grandes chefs puede incluir la frase «cocer a fuego lento el pescado en un vino blanco apropiado hasta que esté casi hecho», pero un algoritmo para el mismo proceso debe comenzar «Escoja un vino blanco cuya etiqueta diga “seco”; coja un sacacorchos y abra la botella; vierta un chorrito de vino en el fondo de la sartén; a fuego fuerte…», es decir, una tediosa división del proceso en pasos muy simples que no requieren decisiones sabias o juicios delicados o intuiciones por parte del lector de la receta.

3) *Garantía de los resultados*: haga lo que haga un algoritmo, siempre lo hace si es ejecutado sin error. Un algoritmo es una receta fácil.

Es sencillo comprobar cómo estas características de los algoritmos hacen posible el ordenador. En el fondo, *todo programa de ordenador es un algoritmo*, compuesto de pasos simples que pueden ser ejecutados con extraordinaria fiabilidad por uno u otro mecanismo simple. El circuito electrónico es habitualmente el mecanismo escogido, pero el poder de los ordenadores no debe nada (salvo la velocidad) a las peculiaridades causales de los electrones moviéndose rápidamente alrededor de los chips de silicona. Los mismos algoritmos pueden ser realizados, aún más rápidamente, mediante dispositivos que derivan fotones en fibra de vidrio y por equipos humanos que utilizan papel y lápiz (¡lo que es mucho, pero que mucho más lento!). La capacidad de los ordenadores para desarrollar algoritmos con tremenda velocidad y fiabilidad ha permitido a los teóricos explorar la peligrosa idea de Darwin por caminos hasta ahora imposibles y con resultados fascinantes.

Lo que Darwin descubrió no era realmente *un* algoritmo, sino más bien una amplia serie de algoritmos relacionados entre sí, de los que él no tenía claro cuál era el método para distinguirlos. Nosotros podemos ahora reformular su idea fundamental de la siguiente manera: «La vida en la Tierra ha sido generada durante billones de años en un único árbol —el árbol de la vida— que se ramifica por uno u otro proceso algorítmico».

Lo que esta formulación significa se hará más claro gradualmente, cuando recorramos las distintas vías por las que la gente ha tratado de expresarlo. En algunas versiones es absolutamente vacío y nada informativo; en otras es manifiestamente falso. Entre medio se encuentran las versiones que realmente

explican el origen de las especies y prometen explicar al mismo tiempo otras cosas. Estas versiones se han ido clarificando con el tiempo gracias a las decididas críticas de aquellos que odian francamente la idea de la evolución como un algoritmo, así como a las impugnaciones de aquellos que la aman.

5. *Los procesos considerados como algoritmos*

Cuando los teóricos piensan en algoritmos, tienen a menudo en mente tipos de algoritmos con propiedades que *no* son compartidas por los algoritmos que nos interesan. Cuando los matemáticos piensan en algoritmos, por ejemplo, generalmente tienen en mente algoritmos que sirven para computar determinadas funciones matemáticas de interés en su trabajo. (La división llamada larga es un ejemplo sencillo. Un procedimiento para ir dividiendo un número elevado entre sus números primos atrae la atención en el exótico mundo de la criptografía). Pero los algoritmos que nos interesan no tienen nada que ver con el sistema de los números u otros objetos matemáticos; son algoritmos para clasificar, seleccionar y construir cosas[13].

Debido a que la mayor parte de las discusiones matemáticas sobre algoritmos tratan de su garantía o de sus comprobables poderes matemáticos, se comete a veces el error elemental de pensar que un proceso que hace uso de la suerte o la aleatorización no es un algoritmos. Pero ¡si incluso las divisiones llamadas largas hacen buen uso de la aleatorización!

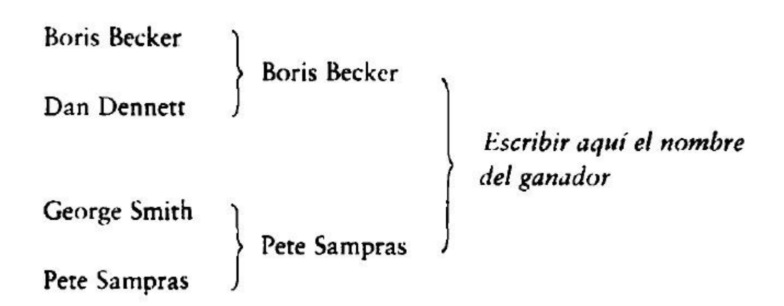
7?

47)326574

¿Cabe seis o siete u ocho veces el divisor en el dividendo? ¿Quién lo sabe? ¿A quién le importa? No lo sabemos; no tenemos ingenio o discernimiento para hacer una división larga. El algoritmo nos lleva directamente a escoger un dígito —al azar, si se quiere— y comprobar el resultado. Si el número escogido vuelve a ser demasiado pequeño, se aumenta en uno y a comenzar de nuevo; si es demasiado largo, se disminuye. Lo bueno de la división larga es que termina funcionando, incluso si uno fuera extraordinariamente estúpido al hacer la primera elección, en cuyo caso se tardaría algo más. Conseguir éxito en tareas difíciles a pesar de una completa estupidez es lo que hace que un ordenador aparezca como algo mágico: ¿cómo puede algo

insensato como una máquina hacer algo tan inteligente? No debe sorprender, por lo tanto, que la táctica de la astuta ignorancia, generando aleatoriamente un candidato para someterlo a prueba mecánicamente, sea un hecho ubicuo en los interesantes algoritmos. No sólo esta característica no interfiere con sus probables poderes como algoritmos sino que es, a menudo, la clave de su poder. (Véase Dennett 1984:149-152, sobre los particularmente interesantes poderes de los algoritmos randomizados de Michael Rabin).

Podemos empezar de cero en el *phylum* de los algoritmos de la evolución considerando los algoritmos cotidianos que comparten con ellos importantes propiedades. Darwin llamó nuestra atención sobre las repetidas oleadas de competición y selección, así que cabe considerar el algoritmo estándar para organizar un torneo eliminatorio, como un torneo de tenis, que eventualmente culmina con los cuartos de final, semifinales y la final, hasta llegar al solitario ganador.



Nótese que este procedimiento reúne las tres condiciones. Es el mismo procedimiento tanto si se dibuja con tiza en una pizarra, se actualiza en el archivo de un ordenador o —¡extraña posibilidad!— no se escribe en ninguna parte sino que simplemente se hace cumplir mediante la construcción de un inmenso abanico de pistas de tenis valladas, cada una con dos puertas de entrada y una sola puerta de salida, la que da paso al ganador a la pista donde se ha de jugar el siguiente partido. (A los perdedores se les dispara un tiro y se les entierra donde han caído). No hay que ser un genio para seguir a los contendientes a lo largo de la prueba, rellenando los espacios en blanco al final de cada partido (o identificando y disparando a los perdedores). Este procedimiento siempre funciona.

Pero ¿qué es lo que *hace* exactamente este algoritmo? Acepta como *input* una serie de competidores y garantiza que terminará identificando un único ganador. Pero ¿qué es un ganador? Depende de la competición. Supongamos que el torneo en cuestión no es de tenis sino de lanzar al aire una moneda a

cara o cruz. Un jugador lanza y el otro elige cara o cruz; el ganador avanza. El ganador del torneo será aquel jugador que haya ganado *n* veces consecutivos lanzamientos de moneda sin perder, dependiendo de cuántas vueltas hayan sido necesarias para completar el torneo.

Hay algo extraño y trivial en este torneo; ¿qué es? El ganador posee una más que notable propiedad. ¿Tenéis noticia de alguien que, a menudo, lanzando la moneda al aire a cara o cruz gane diez veces consecutivas sin perder ninguna? Probablemente nunca ha sucedido. Las probabilidades contra la existencia de tal persona pueden parecer enormes y, en el curso normal de las cosas, seguramente lo son. Si algún apostador os ofrece una probabilidad de 10 a 1 de que él pueda encontrar a alguien que ante vuestros ojos consiga ganar diez veces consecutivas lanzando una moneda no trucada a cara o cruz, podéis inclinaros a pensar que es una buena apuesta. Si es así, cabe esperar que el apostador no consiga 1.024 cómplices (que no sean timadores, sino que jueguen limpio y de buena fe). Esto es lo que hace falta (2^{10} competidores) para organizar un torneo de diez rondas. Al comenzar el torneo, el jugador no tendrá una pista sobre quién será el vencedor, lo que le garantizaría ganar la apuesta, aunque, con toda seguridad, el algoritmo del torneo dará a conocer tal persona en poco tiempo; es una apuesta muy atractiva, con una ganancia asegurada para el jugador. (No me hago responsable de los posibles perjuicios que puede sufrir quien intente hacerse rico aplicando esta golosina de filosofía práctica).

Todo torneo de eliminación termina con un ganador, quien «automáticamente» adquiere cualquiera de las propiedades exigidas para avanzar a lo largo de las eliminatorias, pero, como el torneo de lanzamiento de monedas a cara o cruz demuestra, la propiedad en cuestión *puede* ser no «simplemente histórica»; un hecho trivial acerca de la pasada historia del competidor que no tiene influencia, en último término, sobre sus futuros proyectos. Supongamos, por ejemplo, que las Naciones Unidas decidieran que todos los futuros conflictos internacionales serían dirimidos tirando una moneda al aire, para lo cual cada nación enviaría un representante (si en el conflicto estuviera implicada más de una nación, se organizaría una especie de torneo, cuyas reglas serían las del llamado «round robin», que es un algoritmo diferente). ¿Quién debería ser designado como nuestro representante nacional? Sin duda, el mejor lanzador mundial de monedas. Supongamos que se organiza un enorme torneo eliminatorio para todos los hombres, mujeres y niños en Estados Unidos. Alguien tendría que ganar y quien fuese el ganador ¡debería justamente ganar veintiocho lanzamientos de

monedas consecutivos, sin un fallo! Esto sería un irrefutable hecho histórico para esa persona, pero dado que consideramos el lanzamiento de monedas a cara o cruz como un asunto de suerte, no hay absolutamente razón para creer que el ganador de tal torneo lo haría mejor en una competición internacional que algún otro que fue eliminado en las primeras fases del torneo. La suerte no tiene memoria. Una persona que posee el billete de lotería ganador ha *sido* ciertamente afortunada, y gracias a los millones que ha ganado no tiene necesidad de ser feliz de nuevo, lo cual está bien, dado que es razonable creer que tiene más posibilidades que otra persona para ganar la lotería una segunda vez o para ganar en el siguiente lanzamiento de monedas a cara o cruz en el que participe. (El no darse cuenta de que la suerte no tiene memoria se conoce como la falacia del jugador; es sorprendentemente popular, tan popular que debo subrayar que *es* una falacia, sin ninguna duda o controversia).

En contraste con los torneos de pura suerte, como los lanzamientos de monedas a cara o cruz, hay torneos de habilidades como los de tenis. En este tipo de torneos es razonable creer que los jugadores que alcanzan las últimas rondas lo harían mejor de nuevo si vuelven a jugar con aquellos que perdieron en las primeras rondas. Hay razón para creerlo, pero no es una garantía de que el ganador de este torneo es el mejor de todos, no solamente hoy sino mañana. Con todo, aunque un torneo eliminatorio bien llevado garantiza la proclamación de un ganador, no hay garantía de que un torneo de habilidades identifique realmente al mejor jugador. Esta es la razón por la que, a veces, se dice en la ceremonia inaugural; «Esperemos que gane el mejor», ya que el procedimiento no lo garantiza. El mejor jugador, es decir, el que se considera mejor según los estándares de «ingeniería» de un jugador de tenis (tiene el revés más fiable, el servicio más rápido, mayor resistencia, etc.), puede tener un mal día, una distensión en el tobillo o caer herido por un rayo. Entonces, de manera imprevista, puede ser vencido en la competición por un jugador que no es realmente tan bueno como él. Pero nadie se molestaría organizando o participando en torneos de habilidades si no supiera que, *a la larga*, estos torneos son ganados por los mejores jugadores. *Este hecho* está garantizado por la propia definición de un torneo de habilidades que sea honesto; si no hubiera una probabilidad superior al 50 por ciento de que los mejores jugadores ganen cada ronda, sería un torneo de suerte, no de habilidades.

Habilidad y suerte se mezclan, de modo natural e inevitable, en cualquier competición real, pero las proporciones de esta mezcla pueden variar ampliamente. Un torneo de tenis jugado en pistas de superficie irregular incrementa la proporción del factor suerte, como sucedería si se aplicara una

innovación que consistiese en que a los jugadores, una vez jugado el primer set, y antes de continuar, se les obligara a jugar a la ruleta rusa con un revólver cargado. Pero incluso en este contexto cargado de factor suerte muchos de los mejores jugadores *tenderían*, estadísticamente, a clasificarse para las últimas rondas. El poder del torneo para «discriminar» a largo plazo las diferencias en habilidades puede disminuirse por una catástrofe inesperada aunque, en general, no se reduce a cero. Este hecho, que es tan verdadero en los algoritmos de la evolución como en los de las eliminatorias en los torneos deportivos, es, a veces, pasado por alto por aquellos que hablan o escriben sobre la evolución.

La habilidad, al contrario de la suerte, es *proyectable*; ya que, en las mismas o similares circunstancias, debe ser tenida en cuenta a la hora de repetir actuaciones. Esta relación con las circunstancias pone de manifiesto otro camino por el cual un torneo puede terminar con extraños resultados. ¿Qué pasaría si las condiciones de la competición se cambiaran (como el juego de croquet de *Alicia en el país de las maravillas*)? Si se jugara al tenis en la primera eliminatoria, al ajedrez en la segunda, al golf en la tercera y al billar en la cuarta, no hay razón para suponer que el eventual ganador deba ser particularmente bueno, comparado con la totalidad de los participantes, en *cualquiera* de estas actuaciones; todos los buenos jugadores de golf pueden perder en la eliminatoria de ajedrez y no tener la oportunidad de demostrar su pericia, y si la suerte no desempeña ningún papel en la cuarta eliminatoria con el billar, el ganador puede llegar a ser el segundo *peor* jugador de billar del torneo. En consecuencia, debe haber alguna uniformidad en las condiciones de la competición para que se pueda producir un resultado *interesante* en un torneo.

Pero ¿un torneo —o un algoritmo— ha de producir algo interesante? No. Los algoritmos de los que hablamos casi siempre hacen algo interesante y por esto atraen nuestra atención. Pero un procedimiento no falla por ser un algoritmo sino porque su utilización no es imaginable o valiosa para alguien. Consideremos la posibilidad de una variación en el algoritmo del torneo de eliminación, según la cual los perdedores de las semifinales juegan las finales. Esta es una regla estúpida, que destruye el objetivo de todo el torneo, aunque éste seguirá siendo un algoritmo. Los algoritmos no tienen objetivos ni propósitos. Además de todos los algoritmos utilizados para alfabetizar listas de palabras, hay multitud de algoritmos para alfabetizar erróneamente palabras y funcionan a la perfección (como si nada les importara). También hay algoritmos para hallar la raíz cuadrada de cualquier número excepto el 18

y el 703. Algunos algoritmos hacen cosas tan aburridamente extrañas e inútiles que no se puede decir sucintamente para qué sirven. Simplemente hacen eso y lo hacen continuamente.

Ha llegado el momento de exponer, quizá, la más frecuente incomprensión del darwinismo: la idea de que Darwin demostró la evolución por selección natural como un procedimiento *para producir a los seres humanos*. Desde que Darwin propuso su teoría, la gente ha tratado engañosamente de interpretar esta teoría como la demostración de que está destinada a *nosotros*, de que somos su objetivo, hasta el punto de que la selección y la competición, así como nuestra entrada en escena, estaban garantizadas por la mera participación en el torneo. Esta confusión ha sido fomentada tanto por los amigos como por los enemigos de la evolución, y es análoga a la confusión del vencedor de los torneos de lanzamientos de moneda a cara o cruz, que se complace en la dudosa gloria derivada de su idea de que, dado que el torneo ha tenido un ganador, y dado que él es el ganador, el torneo tenía que terminar proclamándolo a él. La evolución puede ser un algoritmo y la evolución puede habernos producido por un proceso algorítmico, de lo que no se deduce como verdad que la evolución sea un algoritmo para producirnos a nosotros. La principal conclusión de Stephen Jay Gould en su libro *La vida maravillosa* (1989) es que «si volviéramos a rebobinar la cinta de la vida» y la proyectáramos de nuevo, una y otra vez, la posibilidad de que fuéramos *nosotros* el producto final a través del molino de la evolución es infinitesimal. Esto es indudablemente cierto (si entendemos por *nosotros* la particular variedad de *Homo sapiens* que somos: lampiño y erecto, con cinco dedos en cada mano, hablando inglés y francés y jugando al tenis y al ajedrez). La evolución no es un proceso que fue concebido para producirnos, pero de esto no se sigue que la evolución no sea un proceso algorítmico que de hecho nos ha producido. (En el capítulo 10 analizaremos este tema con más detalle).

Los algoritmos de la evolución son sin duda algoritmos interesantes —interesantes, al menos, para nosotros—, no porque lo que garantizan hacer es interesante para nosotros, sino porque hay garantías de que lo que *tienden* a hacer es interesante para nosotros. Los algoritmos son como los torneos de habilidades. El poder de un algoritmo para generar algo de interés o de valor no está limitado a lo que el algoritmo pueda generar de un modo irrefutable y matemáticamente probado, lo que es especialmente cierto en los algoritmos evolutivos. La mayor parte de las controversias sobre el darwinismo, como veremos, se reducen a desacuerdos acerca de hasta qué punto son poderosos

ciertos procesos evolutivos que han sido postulados: ¿pueden hacer realmente todo esto o todo esto en el tiempo disponible? Se trata de investigaciones sobre lo que un algoritmo evolutivo *puede* producir o *podría* producir o *sería posible* que produjera, y sólo indirectamente sobre lo que tal algoritmo *inevitablemente* produciría. El propio Darwin preparó el escenario en el texto de su resumen: su idea es una llamada de atención a lo que con «seguridad» el proceso de la selección natural «tiende» a producir.

Todos los algoritmos están garantizados para hacer cualquier cosa de las que hacen, pero no es necesario que sea algo interesante; algunos algoritmos tienen la garantía de una tendencia (con probabilidad *p*) a hacer algo, cosa que puede o no puede ser interesante. Pero si un algoritmo garantiza hacer algo que no tiene que ser interesante en ningún sentido, ¿cómo vamos a distinguir los algoritmos de otros procesos? ¿Puede cualquier proceso ser un algoritmo? ¿Es el oleaje en una playa un proceso algorítmico? ¿Es el sol, endureciendo la arcilla en el lecho de arena del río reseco, un proceso algorítmico? La respuesta es que ¡hay muchos hechos en estos procesos que *son* mejor apreciados si los consideramos como algoritmos! Analicemos, por ejemplo, la cuestión de por qué los granos de arena de una playa son tan uniformes en el tamaño. Esto es debido al proceso natural que ocurre gracias al repetido lanzamiento de los granos por el oleaje; orden alfabético en gran escala, se podría decir. El dibujo de las grietas que aparecen en la arcilla resecada al sol puede ser mejor explicado si observamos la cadena de acontecimientos que no son distintos de las rondas de un torneo.

Exploremos ahora el proceso del templado de una pieza de metal. ¿Qué proceso puede ser más físico, menos «computacional», que este templado? El herrero calienta repetidamente el metal y luego permite que se enfríe hasta que, en algún momento del proceso, llega a ser más fuerte. ¿Cómo? ¿Qué clase de explicación puede darse a esta mágica transformación? ¿Ha creado el calor átomos de una especial dureza que recubren la superficie? ¿O se trata de un pegamento subatómico aspirado de la atmósfera que aglutina a todos los átomos de hierro? No, no ha sucedido nada parecido a esto. El nivel correcto de la explicación es el nivel algorítmico: cuando el metal se enfría desde su estado fundido, la solidificación comienza en muchos puntos diferentes al mismo tiempo, originando cristales que crecen juntos hasta que todo se solidifica. Pero la primera vez que esto sucede la disposición de estructuras cristalinas individuales es subóptima, mantenidas juntas débilmente y con bastantes tensiones internas y deformaciones. Si se calienta de nuevo, pero sin llegar a ser fundido, se rompen parcialmente estas estructuras, de modo que,

cuando se le permite enfriarse la siguiente vez, los fragmentos rotos se adhieren a los sólidos en una disposición diferente. Se puede probar matemáticamente que estas readaptaciones de la estructura tienden a mejorar progresivamente, aproximándose a la estructura óptima o más fuerte, siempre que el régimen de calentamiento y enfriamiento tenga los parámetros apropiados. Tan poderoso es este procedimiento de optimización que ha sido utilizado como inspiración para una técnica de solución de problemas en la ciencia de la computación: el templado simulado, que no tiene nada que ver con los metales o el calor, pero que es la vía para conseguir un programa de ordenador que sirva para construir, desarmar y volver a construir una estructura de datos (otro programa) una y otra vez, tratando a ciegas de encontrar la mejor u óptima versión (Kirkpatrick, Gelatt y Vecchi 1983). Este ha sido uno de los mayores avances en el desarrollo de la «máquinas Bolzmann» y las «redes Hopfield» y otros esquemas que permiten resolver a satisfacción las limitaciones en los procesos algorítmicos, que son las bases para los trabajos de los modelos conectistas o arquitecturas de redes nerviosas en la inteligencia artificial. (Para revisiones de estos problemas, véanse Smolensky 1983, Rumelhart 1989, Churchland y Sejnowski 1992 y, a nivel filosófico, Dennett 1987a y Paul Churland 1989).

Si se desea profundizar en el conocimiento de cómo funciona el templado en metalurgia debemos aprender naturalmente la física de todas las fuerzas que operan a nivel atómico, aunque debe subrayarse que la idea básica de cómo funciona el templado (y especialmente *por qué funciona*) puede ser conseguida libre de aquellos detalles; después de todo, yo lo he explicado en términos simples, no de expertos (¡y no conozco la física!). La explicación del templado puede ser realizada con una terminología básica neutra: debemos esperar que una optimización de un cierto tipo ocurra en cualquier «material» cuyos componentes han sido dispuestos juntos mediante cierto tipo de proceso de construcción y que puede ser descompuesto en unas secuencias con el objetivo de modificar un único parámetro global, etc. Esto es lo que es común para todos los procesos que van desde la barra de acero caliente hasta el poderoso superordenador.

Las ideas de Darwin sobre el poder de la selección natural pueden ser también separadas de su base biológica. Efectivamente, como ya hemos señalado, el propio Darwin poseía pocos indicios (y los que tenía los había desestimado por erróneos) acerca de cómo se realizaban los procesos microscópicos de la herencia genética. Aunque no conocía los detalles de su sustrato físico, no obstante Darwin podía discernir que si se reunían ciertas

condiciones se fraguarían ciertos efectos. Esta neutralidad del sustrato había sido crucial al permitir que las ideas básicas de Darwin flotaran como un corcho en las olas de las subsiguientes investigaciones y controversias, y esto hizo que lo que había sucedido desde Darwin adquiriese un curioso movimiento de vaivén. Darwin, como hemos señalado en el capítulo precedente, nunca atinó con la absolutamente necesaria idea de un gen, aunque, de modo paralelo, llegó el concepto de Mendel para poner a la mano precisamente la correcta estructura que pudiera dar un sentido matemático a la herencia (y resolviendo el desagradable problema de Darwin con respecto a la mezcla de los caracteres hereditarios). Y entonces, cuando fue identificado el ADN como el real vehículo físico de los genes, pareció a primera vista (y aún aparece así para muchos participantes en las discusiones) como si los genes de Mendel pudieran ser simplemente *identificados* como un fragmento particular de ADN. Pero las cosas aparecían como mucho más complejas; mientras más aprendían los científicos sobre la real biología molecular del ADN y su papel en la reproducción, se hacía más claro que la historia de Mendel era, como mucho, una enorme hipersimpliíicación. ¡Alguno fue tan lejos como para decir que hemos aprendido recientemente *que no hay genes mendelianos*! Habiendo subido la escalera de Mendel, debemos ahora abandonarla. Pero naturalmente nadie desea abandonar una herramienta tan valiosa, que aún se prueba a sí misma diariamente en cientos de contextos científicos y técnicos. La solución es subir a Mendel de nivel y declarar que él, como Darwin, captó una verdad *abstracta* acerca de la herencia. Nosotros podemos, si queremos, hablar de *genes virtuales*, considerando que estos genes tienen su realidad distribuida en el concreto material del ADN. (Queda bastante por decir en favor de esta opción, la cual será discutida más ampliamente en los capítulos 5 y 12).

Pero ahora, retornando a la cuestión planteada más arriba, ¿existen límites a lo que puede ser considerado un proceso algorítmico? Yo creo que la respuesta es no; cualquier proceso a nivel abstracto, si se desea, puede tratarse como un proceso algorítmico. ¿Qué significa esto? En realidad, sólo algunos procesos, cuando son tratados como algoritmos, generan resultados interesantes, aunque no tenemos que definir lo que es un «algoritmo» de tal modo que incluya únicamente a los interesantes (¡un orden filosófico excesivo!). El problema se resuelve solo, ya que nadie perderá el tiempo examinando los algoritmos que no son interesantes, por una u otra razón. Todo depende de lo que necesitemos explicar. Si lo que nos sorprende como un rompecabezas es la uniformidad de los granos de arena o la dureza de la

hoja de una espada, una explicación algorítmica es la que podrá satisfacer nuestra curiosidad; esta será la verdad. Otros interesantes hechos del mismo fenómeno o de los procesos que lo han originado, pueden no adaptarse a un tratamiento algorítmico.

Aquí está, pues, la peligrosa idea de Darwin: el nivel algorítmico es el nivel que mejor explica la velocidad del antílope, las alas del águila, la forma de la orquídea, la diversidad de las especies y todas aquellas ocasiones en las que el mundo de la naturaleza nos maravilla. Es difícil creer que algo tan estúpido y tan mecánico como un algoritmo pueda producir cosas tan admirables. Por muy impresionantes que sean los productos de un algoritmo, siempre el proceso subyacente no consiste más que en un conjunto de pasos individuales, no inteligentes, que se suceden uno tras otro sin la ayuda de una supervisión inteligente; por definición, son «automáticos»: son los trabajos de un autómata. Se alimentan unos a otros o por el ciego azar —si se quiere, dependiendo del lado que cae la moneda— y nada más. La mayoría de los algoritmos que nos son familiares originan los productos más modestos: hacen divisiones largas o listas alfabéticas o la declaración de la renta del contribuyente medio. Los algoritmos más imaginativos producen los gráficos animados creados por ordenador que vemos cada día en la televisión, transformando caras, creando hordas de osos polares imaginarios que patinan sobre el hielo, simulando mundos totalmente virtuales de entes nunca vistos o imaginados. Pero la biosfera real es mucho más fascinante, en muchos ordenes de magnitud. ¿Puede ser realmente esta biosfera el resultado de nada más que una cascada de procesos algorítmicos alimentados por la suerte? Y si es así, ¿quién diseñó la cascada? Nadie. Es el producto de un ciego proceso algorín mico. Como el propio Darwin señaló en una carta al geólogo Charles Lyell poco después de la publicación de *El origen de las especies*: «Yo no daría absolutamente nada por la teoría de la selección natural si requiere adiciones milagrosas en cualquiera fase del descenso evolutivo… Si yo estuviera convencido de la necesidad de tales adiciones para la teoría de la selección natural, la rechazaría como basura» (Darwin 1911:II, 6-7).

De acuerdo con Darwin la evolución es un proceso algorítmico. Dicho de esta manera, todavía es motivo de controversia. Una de las guerras de prestigio que están en marcha dentro de la biología evolucionista es la entablada entre aquellos que continuamente presionan y presionan hacia un tratamiento algorítmico y aquellos que, por varias razones ocultas, se resisten a esta tendencia. Es como si se tratara de metalúrgicos descontentos con la explicación algorítmica del templado. «¿Quiere decirme que el proceso del

templado es así? ¿No es necesario un superpegamento submicroscópico especialmente creado por el proceso de calentamiento y enfriamiento?». Darwin ha convencido a todos los científicos que la evolución, como el templado, *funciona*. Su visión radical de *cómo y por qué* funciona es aún combatida, sobre todo porque aquellos que resisten pueden entrever que sus escaramuzas forman parte de una larga campaña. Si se pierde el juego de la biología evolucionista, ¿dónde terminará todo?

CAPÍTULO 3
El ácido universal

I. *Primeras reacciones*

> Ahora, el origen del hombre ha sido demostrado… La metafísica debe florecer… Quien entendiera al babuino haría más por la metafísica que el propio Locke.
>
> *Charles Darwin*
> (Citado en P. H. Barren y otros, *Charles Darwin's Notebooks*)

> Su tema es el «origen de las especies» y no el origen de la organización; parece una mala jugada innecesaria haber iniciado esta última especulación.
>
> *Harriet Martineau*
> (Carra a Darwin de 13 de marzo de 1860, citada en Desmond y Moore, *Darwin*)

Darwin inició su explicación desde la zona media del proceso evolutivo e incluso podría decirse que desde su final: comenzando con las formas de vida que observamos en el presente y demostrando cómo los modelos de la biosfera de nuestro tiempo pueden ser explicados como consecuencias del proceso de la selección natural, a partir de modelos previos de la biosfera de ayer y así, retrocediendo, hasta el más distante pasado. Darwin comenzó con hechos que todo el mundo conoce: todos los seres vivos actuales son hijos de padres que son hijos de abuelos y así sucesivamente, de modo que todo lo que es vivo hoy es una rama de una familia genealógica, la cual, a su vez, es una rama de un clan mucho más grande. Darwin seguía argumentando que si retrocedemos lo suficientemente lejos, nos encontraremos con que todas las ramas de todas las familias eventualmente nacen de comunes troncos ancestrales, así que hay un único árbol de la vida, con todas sus grandes ramificaciones, ramas y ramillas unidas mediante el denominado, en la teoría de la evolución, «descenso con modificación». El hecho de que exista una organización arbórea es crucial para explicar el tipo de proceso implicado, porque el árbol podría haber sido creado por un proceso automático y recursivo: primero se construye una *x*, entonces se modifican los descendiente de *x*, después se modifican aquellas modificaciones y después se modifican

las modificaciones de las modificaciones… Si la vida es un árbol, todo podría haber aparecido mediante un proceso inexorable, automático y remodelador en el cual los diseños se acumularían a lo largo del tiempo.

Operando en sentido *retrógrado*, comenzando al «final» o cerca del final de un proceso, la resolución del penúltimo paso antes de interrogarnos cómo ha podido ser producido *esto*, es un método seguro y fiable utilizado por quienes diseñan programas de ordenadores, especialmente cuando utilizan el método conocido como «recursion». Habitualmente es una cuestión de humildad práctica: si no se desea morder más de lo que se puede masticar, el bocado correcto para comenzar es, a menudo, el último bocado, si se puede encontrar. Darwin lo encontró y entonces trabajó muy cuidadosamente hacia atrás, soslayando muchos de los grandes temas que su investigación había agitado, meditando sobre ellos en sus cuadernos de notas personales, aunque posponiendo la publicación de estos cuadernos indefinidamente. (Por ejemplo, Darwin deliberadamente evitó discutir la evolución humana en *El origen de las especies*; véase la discusión en R. J. Richards 1987:160 ss.). Pero Darwin, como muchos de sus lectores, supo ver adonde conducía todo esto, a pesar de su casi perfecto silencio sobre estas perturbadoras extrapolaciones. A algunos les gustó lo que ellos pensaban que habían entrevisto y otros lo odiaron.

Karl Marx estaba exultante: «No sólo es un golpe mortal por primera vez para la “Teleología” en las ciencias naturales, sino que su significado racional es explicado empíricamente» (citado por Rachels 1991:110). Friedrich Nietzsche vislumbró —a pesar de su desprecio por todo lo inglés— un mensaje incluso más cósmico en la idea de Darwin: Dios ha muerto. Si Nietzsche es el padre del existencialismo, es posible que Darwin merezca el título de abuelo. Otros quedaron menos subyugados con la creencia de que las ideas de Darwin eran absolutamente subversivas para la tradición sagrada. Samuel Wilberforce, obispo de Oxford, cuyo debate con Thomas Huxley en junio de 1860 fue uno de los más celebrados enfrentamientos entre los darwinistas y el *establishment* religioso (véase el capítulo 12), escribió en una reseña anónima:

> La supremacía del hombre sobre la tierra; su capacidad de articular el lenguaje; el don de la razón; su libre albedrío y su responsabilidad… —todas estas cosas son por igual absolutamente irreconciliables con la degradante noción del origen animal de quien fue creado a imagen de Dios… (Wilberforce 1860).

Cuando surgieron las especulaciones sobre las implicaciones de su idea, Darwin sabiamente escogió retirarse a la seguridad de su campo base, es

decir, la magníficamente aprovisionada y bien defendida tesis que comenzó en el medio del proceso evolutivo, con la vida ya en escena, y mostrando «simplemente» cómo, una vez que el proceso de acumulación de diseños estaba en marcha, podía seguir adelante sin ninguna (¿otra?) intervención de Mente alguna. Pero, como muchos de sus lectores apreciaron, por muy confortante que esta modesta renuncia pudiera ser, no era realmente un lugar de descanso estable.

¿Han oído hablar del ácido universal? Esta fantasía era motivo de diversión para mí y para algunos de mis compañeros de colegio; no sé si la inventamos o la heredamos, junto con la *Spanish fly* y el salitre, como una parte de nuestra juvenil cultura underground. ¡El ácido universal es un líquido tan corrosivo que atravesará *cualquier cosa*! El problema es: ¿dónde guardarlo? Disuelve las botellas de cristal y los recipientes de acero inoxidable con tanta facilidad como las bolsas de papel. ¿Qué sucedería si alguien descubriera o llegara a descubrir o crear una masa de ácido universal? ¿Sería eventualmente destruida la totalidad del planeta? ¿Qué quedaría como resultado de su acción? Después de que todo haya sido transformado tras su encuentro con el ácido universal, ¿qué aspecto tendría el mundo? No me imaginé que en pocos años me encontraría una idea —la idea de Darwin— con un indudable parecido con el ácido universal: corroe prácticamente todos los conceptos tradicionales y deja como resultado una visión revolucionaria del mundo, con la mayoría de sus viejos hitos todavía reconocibles pero transformados sustancialmente.

La idea de Darwin había nacido como respuesta a preguntas en el ámbito de la biología, pero amenazaba con extenderse, aportando respuestas —bienvenidas o no— a interrogantes que afectaban a la cosmología (en una dirección) y a la psicología (en la otra dirección). Si el *re*diseño puede ser un proceso algorítmico de evolución, no inteligente, ¿por qué no podía ser todo el proceso un producto de la evolución, y así hasta el principio de los tiempos? Y si la evolución no inteligente puede ser responsable de los soberbios e inteligentes artefactos de la biosfera, ¿cómo podían quedar exentos de una explicación evolutiva los productos de nuestras propias «reales» mentes? La idea de Darwin, por lo tanto, también amenazaba con extenderse hasta el final de los tiempos, disolviendo la ilusión de nuestra propia autoría, nuestra divina chispa de creatividad y entendimiento.

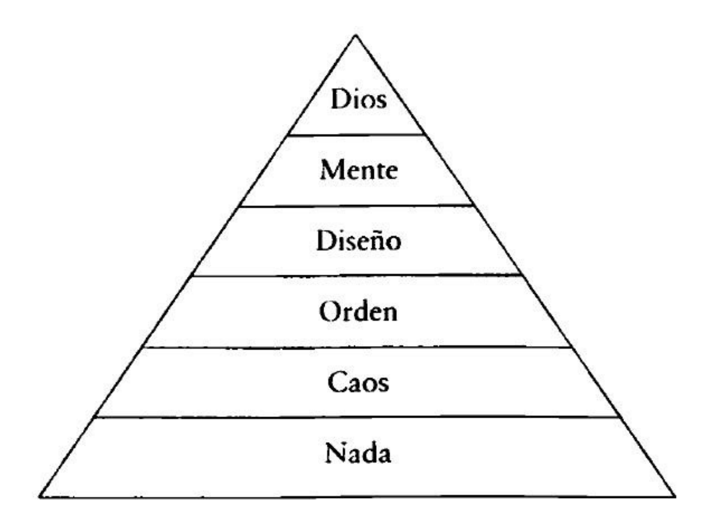
Una buena parte de las controversias y ansiedades que han envuelto a la idea de Darwin desde entonces puede ser comprendida incluso como una serie de campañas fallidas en la lucha por mantener la idea de Darwin dentro de

una revolución parcial, y aceptablemente «segura». ¡Quizá ceder a Darwin alguna parte o la totalidad de la moderna biología, pero pararle los pies ahí! ¡Mantener el pensamiento darwiniano fuera de la cosmología, fuera de la psicología, fuera de la cultura humana, fuera de la ética, la política y la religión! En estas campañas muchas batallas han sido ganadas por las fuerzas de contención: aplicaciones erróneas de la idea de Darwin han sido expuestas y desacreditadas, batidas en retirada por los paladines de la tradición predarwiniana. Pero siguen llegando nuevas oleadas de pensamiento darwiniano. Parecen versiones mejoradas, más difíciles de refutar que sus predecesoras, pero ¿son ramificaciones del sólido núcleo de la teoría de Darwin, o quizá son perversiones de ella, incluso más virulentas y más peligrosas que las tergiversaciones de Darwin ya refutadas anteriormente?

Los que se oponen a la extensión de la idea de Darwin tienen diferentes opiniones sobre tácticas a seguir. ¿Dónde debemos construir los diques de contención? ¿Debe contenerse la idea dentro de la propia biología, con una u otra contrarrevolución postdarwiniana? Entre aquellos que han favorecido esta táctica se encuentra Stephen Jay Gould, quien ha postulado diferentes revoluciones para frenar la idea de Darwin. ¿O deberíamos colocar las barreras más adelante? Para evaluar nuestras fuerzas en esta serie de campañas debemos comenzar con un simple mapa del territorio predarwiniano. Como veremos, tendrá que ser revisado, una y otra vez, para reacomodarlo, cuando se hayan perdido varias escaramuzas.

2. *El asalto de Darwin a la pirámide cósmica*

Un hecho destacado de las visiones predarwinianas del mundo es un mapa global de las cosas, desde su cúspide hasta su base. A menudo se describe como una escalera; Dios se encuentra en la cúspide, con los seres humanos un peldaño o dos por debajo (dependiendo de si los ángeles forman parte o no del esquema). En la base de la escalera se encuentra la nada o tal vez el caos o tal vez la materia de Locke, inerte e inmóvil. Alternativamente la escalera es una torre o, en la memorable frase del historiador intelectual Arthur Lovejoy [1936], una gran cadena de seres, compuesta por muchos eslabones. El argumento de John Locke ha dirigido ya nuestra atención hacia una particular versión abstracta de la jerarquía, a la que hemos llamado la pirámide cósmica:



(Advertencia: ¡Cada término de la pirámide debe ser entendido en un sentido predarwiniano pasado de moda!).

Todas las cosas encuentran su lugar en uno u otro nivel de la pirámide cósmica, incluso la vacía nada, el último fundamento. No toda la materia está ordenada, alguna se halla en estado de caos; sólo alguna materia ordenada se encuentra también diseñada; sólo algunas cosas diseñadas tienen también mentes y, naturalmente, sólo una mente es Dios. Dios, la primera mente, es la fuente y explicación de todas las cosas por debajo de él. (Dado que todo *depende* de Dios, quizá deberíamos decir que la figura sería más un candelabro colgando de Dios, que una pirámide que Le sostuviera).

¿Cuál es la diferencia entre orden y diseño? A primera vista, podemos decir que el orden es mera regularidad, un simple modelo, mientras que el diseño corresponde al *telos* de Aristóteles, como explotación del orden para un propósito, tal como vemos en un artefacto diseñado inteligentemente. El sistema solar exhibe un magnífico orden pero no tiene (aparentemente) un propósito; no es *para* algo. Un ojo, por el contrario, es *para* ver. Antes de Darwin, la distinción no era tan clara, sino positivamente borrosa:

En el siglo XIII, Santo Tomás de Aquino defendía que los cuerpos naturales (como los planetas, las gotas de la lluvia y los volcanes) actuaban como si estuvieran dirigidos hacia un objetivo definido, «para obtener el mejor resultado». Esta adecuación de los medios a los fines implica, argumentó Santo Tomás de Aquino, una intención. Pero, teniendo en cuenta que los cuerpos naturales no tienen conciencia, no pueden proporcionar esta intención por sí mismos. «Por lo tanto existe algún ser inteligente que dirige todas las cosas naturales hacia su fin: y a este ser lo llamamos Dios» (Davies 1992:200).

Siguiendo esta tradición, el Cleantes de Hume incluye las adaptadas maravillas del mundo viviente con las regularidades del cielo; para él *todo* es como un maravilloso mecanismo de relojería. Pero Darwin sugiere una división: Dadme orden, dice, y tiempo y os daré diseño. Permitidme comenzar con regularidad —la mera regularidad de la física, sin propósito, sin inteligencia y sin objetivo— y os mostraré un proceso que eventualmente generará productos que exhiben no sólo precisamente regularidad sino un diseño con propósito. (Esto es justamente lo que Karl Marx pensó cuando declaró que Darwin había lanzado un golpe mortal a la teleología: Darwin había *reducido* la teleología a no teleología, el diseño al orden).

Antes de Darwin, la diferencia entre orden y diseño no parecía grande, debido a que, en cualquier caso, todo procedía de Dios. La totalidad del universo era Su artefacto, un producto de Su inteligencia, de Su mente. Una vez que Darwin se plantó en el medio con su respuesta a la pregunta de cómo el diseño podría surgir del simple orden, el resto de la pirámide cósmica se tambaleó. Supongamos que aceptamos que Darwin explicó el diseño de los cuerpos de las plantas y de los animales (incluyendo nuestros cuerpos, ya que admitimos que Darwin nos había ubicado firmemente en el reino animal). Si miramos hacia arriba, si le concedemos a Darwin nuestros cuerpos, ¿seremos capaces de impedir que se apropie de nuestras mentes? (Trataremos esta cuestión, ampliamente, en la Parte III). Si miramos hacia abajo, Darwin nos pide que/le concediéramos el orden como una premisa, pero ¿hay algo que le impida descender un nivel y darnos una explicación algorítmica del origen del orden a partir del simple caos? (Trataremos esta cuestión en el capítulo 6.)

El vértigo y la repulsa que este proyecto provoca en tantas personas fue perfectamente expresado en un temprano ataque a Darwin publicado anónimamente en 1868:

> En la teoría con la que nos enfrentamos la Absoluta Ignorancia es el artífice; así que podemos enunciar como el principio fundamental de la totalidad del sistema que PARA HACER UNA PERFECTA Y HERMOSA MÁQUINA, NO ES NECESARIO SABER CÓMO HACERLA. En un cuidadoso examen, encontramos que esta proposición expresa en forma condensada el significado esencial de la teoría de Darwin y expresa; en unas pocas palabras, toda la explicación de Mr. Darwin: quien, por una extraña inversión del razonamiento, parece creer que la Ignorancia Absoluta está totalmente cualificada para ocupar el lugar de la Sabiduría Absoluta en todos los logros de la habilidad creativa (Mackenzie 1868).

¡Exactamente! La «extraña inversión del pensamiento» de Darwin fue, de hecho, una nueva y admirable forma de pensar, sobrepasando la vía de la primera mente que John Locke había «demostrado» y que David Hume supo Refutarle. John Dewey describió bellamente esta inversión algunos años más tarde, en su clarividente libro *The Influence of Darwin on Philosophy*: «El

interés se desplaza… desde una inteligencia que moldeó cosas de una vez por todas hacia las particulares inteligencias que aún siguen modelando las cosas» (Dewey 1910:15). Pero la idea de tratar a la Mente como un efecto, más que como una primera causa, es demasiado revolucionaria —algo tremendamente «difícil de creer»— para que algunas mentes puedan asimilarlo cómodamente. Esto es tan cierto hoy como lo era en 1860, y se ha podido aplicar siempre tanto a los mejores amigos de la evolución como a sus enemigos. Por ejemplo, el físico Paul Davies, en su reciente libro *La mente de Dios*, proclama que el poder reflexivo de las mentes humanas puede ser «no un detalle trivial, no un subproducto menor de fuerzas no inteligentes y sin propósito» (Davies 1992:232). Esto es una forma muy reveladora de expresar un rechazo que suena a familiar porque delata un prejuicio nial analizado. ¿Por qué, podríamos preguntar a Davies, no puede la cosa más importante de todas ser algo que surge de cosas no importantes? ¿Por qué la importancia o excelencia de *alguna cosa* debe llegar desde las alturas, desde algo más importante, como un regalo de Dios? La inversión de Darwin sugiere que abandonemos esta presunción y que busquemos esos tipos de excelencia, de valor y propósito, que pueda emerger, desde «fuerzas no inteligentes y sin propósito».

Alfred Russel Wallace, cuya propia versión de la evolución mediante selección natural llegó a la mesa de trabajo de Darwin mientras éste seguía retrasando la publicación de *El origen de las especies*, y a quien Darwin intentó tratar como codescubridor de dicho principio, nunca entendió este problema**[14]**. Aunque al final Wallace era bastante más decidido en el tema de la evolución de la mente humana de lo que Darwin deseaba, y mantenía al principio resueltamente de que la mente humana no era una excepción a la regla de que todos los seres vivos eran productos de la evolución, Wallace no llegó a considerar la «extraña inversión del razonamiento» como la clave de la trascendencia de la gran idea. Haciéndose eco de John Locke, Wallace proclamó que «la maravillosa complejidad de las fuerzas que parecen controlar la materia, si no constituirla realmente, son y deben ser productos de la mente» (Gould 1985:397). Cuando, más tarde, Wallace se convirtió al espiritualismo y dejó exenta a la conciencia humana de la regla de hierro de la evolución, Darwin se dio cuenta de la amplitud de la grieta provocada en su idea y le escribió: «Espero que no haya asesinado completamente a su propia criatura y a la mía» (Desmond y Moore 1991:569).

Pero ¿fue realmente tan inevitable que la idea de Darwin abocara a tal revolución y subversión? «Es obvio que los críticos no tenían voluntad de

comprender y, en alguna medida, el propio Darwin estimuló esta forma de pensar» (Ellegard 1956). Wallace deseaba plantearse cuál podría ser el *propósito* de la selección natural. Aunque pueda parecer visto retrospectivamente como un despilfarro de la fortuna que él y Darwin habían descubierto, era una idea por la cual el propio Darwin había expresado su simpatía. En lugar de reducir la teleología definitivamente a un orden sin propósito, ¿por qué no reducir toda la teleología del mundo a un solo propósito: el propósito de Dios? ¿No era éste un camino obvio y atrayente para taponar la brecha en el dique? Darwin tenía claro que la variación de la que depende el proceso de la evolución natural *había* de ser no planeada y no diseñada, pero ¿debía de tener el propio proceso un propósito? En una carta de 1860 al naturalista norteamericano Asa Gray, uno de sus primeros seguidores, Darwin escribió: «Me inclino a considerar todas las cosas como el resultado de unas leyes diseñadas (con este énfasis), con los detalles, ya sean buenos o malos, dejados en manos del proceso al que podemos llamar azar» (F. Darwin 1911:11, 105).

A menudo, los procesos automáticos son en sí mismos creaciones de gran brillantez. Desde la ventajosa posición actual, podemos ver que los inventores de la transmisión y la apertura automática de las puertas no eran idiotas y su genio radica en haber sabido crear algo que pudiera hacer algo «inteligente» sin tener que pensar en ello. Aceptando algún anacronismo, podríamos decir que para algunos observadores en los días de Darwin parece como si él hubiera dejado abierta la posibilidad de que Dios hiciese Su trabajo manual por designio de un diseñador automático. Para algunos de aquellos observadores, la idea no era precisamente un freno desesperado sino un positivo mejoramiento de la tradición. El primer capítulo del Génesis describe las sucesivas etapas de la Creación y cada una de ellas acaba con la frase «y Dios vio que esto era bueno». Darwin había descubierto un camino para eliminar esta aplicación al por menor del control de calidad inteligente; la selección natural se encargaría de este control sin la intervención divina. (El filósofo del siglo XVII Gottfried Wilhelm Leibniz había defendido una visión parecida a la del Dios el Creador). Como Henry Ward Beecher escribió: «El diseño al por mayor es de mayor grandeza que el diseño al por menor» (Rachels 1991:99). Asa Gray, cautivado por la nueva idea de Darwin, aunque tratando de reconciliarla con su tradicional credo religioso en cuanto fuera posible, siguió adelante con este matrimonio de conveniencia: Dios *propone* la «corriente de variaciones» y prevé de este modo que las leyes de la naturaleza que Él ha formulado recorten esta corriente a lo largo de los eones.

Más tarde, como John Dewey explicó muy bien, mediante otra metáfora mercantil: «Gray mantuvo lo que puede ser llamado diseño sobre el plan de instalación» (Dewey 1910:12).

No es inusual encontrar este tipo de metáforas impregnadas de capitalismo en las explicaciones de la evolución. Los ejemplos son, a menudo, jubilosamente recogidos por aquellos críticos e intérpretes de Darwin que ven este lenguaje como una revelación —o deberíamos decir traición— del ambiente social y político en el que Darwin desarrolló sus ideas, y, de este modo, desacreditando sus llamadas a la objetividad científica. Es verdad que Darwin, como simple mortal, había heredado múltiples conceptos, modos de expresión, actitudes, sesgos y visiones del mundo que iban ligados al tiempo en que vivió (como un inglés Victoriano), pero también es verdad que las metáforas económicas que surgen con tanta naturalidad en la mente cuando uno piensa en la evolución consiguen su poder de uno de los hechos más profundos del descubrimiento de Darwin.

3. *El principio de la acumulación de diseños*

La clave para comprender la contribución de Darwin es la concesión de la premisa del argumento del diseño. ¿Qué conclusión debe extraerse si se encuentra un reloj en una selva? Como Baley insiste (y el Cleantes de Hume antes que él), un reloj muestra una enorme cantidad de *trabajo hecho*. Los relojes y otros objetos diseñados no aparecen por las buenas, sino que son el producto de lo que la industria moderna denomina I+D, es decir, Investigación más Desarrollo, y la I+D es costosa, tanto en tiempo como en energía. Antes de Darwin, el único modelo que teníamos de un proceso mediante el cual esta suerte de trabajo de I+D pudiera ser hecho era un Artífice Inteligente. Lo que Darwin intuyó era que, en principio, el mismo trabajo podía ser realizado por una clase diferente de procesos que *distribuían* este trabajo a lo largo de enormes períodos de tiempo, conservando económicamente el trabajo de diseño que se había conseguido en cada etapa, de modo que no tuviera que ser hecho de nuevo. En otras palabras, Darwin había hecho blanco en el que puede ser llamado el principio de la acumulación de diseños. Las cosas en el mundo (como los relojes y los organismos, entre otras) pueden ser vistas como productos que incorporan una cierta cantidad de diseño y, por un camino o por otro, este diseño debe ser creado por un proceso de I+D. Lo absolutamente no sujeto a designio —el puro caos, en su antiguo sentido— era el cero, el punto de partida.

Una idea más reciente respecto a la diferencia —y estrecha relación— entre diseño y orden nos ayudará a clarificar esta imagen. Es la propuesta protagonizada en primer lugar por el físico Erwin Schrödinger (1967) de que la vida puede ser definida en términos de la segunda ley de la termodinámica. En física, el orden u organización puede ser medido en términos de *diferencias de calor* entre regiones de espacio/tiempo; la entropía es simplemente desorden, lo opuesto al orden, y de acuerdo con la segunda ley, la entropía de un sistema aislado aumenta con el paso del tiempo. En otras palabras, las cosas se van degradando, inevitablemente. De acuerdo con la segunda ley, el universo se va relajando desde un estado más ordenado hasta un estado final completamente desordenado conocido como muerte calórica (*heat death*) en el cual todas las cosas se encuentran a la misma temperatura y la entropía es máxima[15].

¿Qué son, entonces, los seres vivos? Son cosas que resisten su desintegración, al menos durante algún tiempo, por no estar aisladas, captando de su entorno los requisitos para mantener viva su rama en el árbol de la vida. El psicólogo Richard Gregory ha resumido con agudeza esta idea:

> La flecha del tiempo derivada de la entropía —la pérdida de la organización, o pérdida de las diferencias de temperatura— es estadística y es un tema para inversiones locales a pequeña escala. Mucho más sorprendente: la vida es una inversión sistemática de la entropía, y la inteligencia crea estructuras y energías diferentes contra la supuesta «muerte» gradual mediante la entropía del universo físico (Gregory 1981:136).

Gregory otorga a Darwin el crédito de la fundamental y poderosa idea: «Es la medida del concepto de la selección natural la que incrementa en complejidad y orden de los organismos en el tiempo biológico la que ahora puede ser entendida». No solamente los organismos individuales, sino todo el proceso evolutivo que los ha creado, de este modo puede ser visto como fenómenos físicos fundamentales que se desarrollan en sentido contrario al tiempo cósmico, un hecho detectado por William Calvin en uno de los significados del título de su clásica exploración sobre las relaciones entre evolución y cosmología. *The River That Flows Uphill: A Journey from the Big Bang to the Big Brain*, ‘El río que fluye hacia la cima de la montaña: Un viaje desde el Big Bang hasta el Gran Cerebro’ (1986).

Por lo tanto, una cosa que ha sido diseñada es un ser vivo o una parte de un ser vivo o un artefacto de un ser vivo, organizado en cualquier caso como ayuda para la batalla contra el desorden. No es imposible oponerse a la tendencia que marca la segunda ley, pero es muy costoso. Veamos lo que sucede con el hierro. Se trata de un elemento muy útil, esencial para nuestra

salud corporal y también valioso como principal componente del acero, ese admirable material de construcción. Nuestro planeta posee grandes reservas de hierro aunque gradualmente van disminuyendo. ¿Significa esto que la Tierra se está quedando sin hierro? No exactamente. Con la trivial excepción de unas cuantas toneladas que han sido lanzadas recientemente fuera del campo gravitacional efectivo de la Tierra como componentes de sondas espaciales, hay la misma cantidad de hierro que había antes. El problema es que cada vez más cantidad de este metal está disperso en forma de herrumbre (moléculas de óxido de hierro) y otros materiales de baja concentración. En principio podría ser recuperado, pero hacerlo costaría enormes cantidades de energía, aplicadas ingeniosamente en un proyecto concreto de extracción y reconcentración.

La organización de un proceso tan complejo es lo que constituye el sello distintivo de la vida. Gregory dramatiza este hecho con un ejemplo inolvidable. Una expresión habitual en los libros de texto acerca de la dirección impuesta hacia la entropía por la segunda ley de la termodinámica es la afirmación de que un huevo revuelto no se puede *recomponer*. No quiere decir en absoluto que no pueda hacerse, sino de que sería una tarea extremadamente costosa y complicada, a contracorriente de la segunda ley de la termodinámica. ¿Cuanto costaría construir un dispositivo en el que se introdujera un huevo revuelto como *input* y devolviera un huevo intacto, como *output*? Hay una rápida solución: introducir una gallina en una caja y alimentarla con los huevos revueltos. Normalmente, las gallinas no llaman la atención por ser seres complejos y casi milagrosos, pero hay una cosa que la gallina puede hacer gracias al diseño que la ha organizado así y que sobrepasa con mucho los dispositivos creados por los ingenieros humanos.

Mientras más diseño exhibe una cosa, más y más trabajo de 1+D ha sido necesario para producirla. Como cualquier buen revolucionario, Darwin explotó el viejo sistema todo lo que le fue posible: retuvo la dimensión vertical de la pirámide cósmica y la convirtió en la medida de la cantidad de diseño que interviene en los componentes de la pirámide. En el esquema de Darwin, como en la pirámide tradicional, las mentes se encuentran cerca de la cúspide, entre las entidades más diseñadas (en parre porque son cosas autodiseñadas, como veremos en el capítulo 13). Esto significa que se encuentran entre los más avanzados *efectos* (hasta la fecha) del proceso creativo, y no —como en la vieja versión— su causa o fuente. Por otra parte, sus productos —los artefactos humanos que fueron nuestro modelo inicial— deben contarse entre las cosas con diseño. Esto parece ir, a primera vista,

contra nuestra intuición. Una oda de Keats puede parecer que tiene alguna justificación para considerarla con un «pedigrí» mayor en I+D que una mariposa —o al menos así puede parecerle a un poeta ignorante de la biología —, pero ¿qué puede decirse de un *clip* para sujetar papeles? Seguramente un *clip* es, como diseño, un producto sencillo comparado con cualquier ser vivo, por rudimentario que éste sea. En un sentido menos obvio, es así, pero reflexionemos un momento. Calcémonos los zapatos de Paley y caminemos por una playa aparentemente desierta en un extraño planeta. ¿Qué descubrimiento nos excitaría más: una almeja o un rastrillo para recoger almejas? Antes de que el planeta pueda fabricar un rastrillo para almejas, tendría que crear un hacedor de rastrillos para almejas y éste es un ente que acumula de largo más diseños que una almeja.

Sólo una teoría con la estructura lógica de la teoría de Darwin podría *explicar* cómo llegan a existir las cosas con diseño, ya que cualquier otro tipo de explicación sería un círculo vicioso o un infinito retorno (Dennett 1975). La vieja vía, la vía de Locke con su primera mente, acepta el principio de que es necesaria una Inteligencia para crear una inteligencia. Para nuestros antepasados, los hacedores de artefactos —hasta alcanzar retrospectivamente al *Homo habilis*, el hombre dotado de manos, de quien desciende el *Homo sapiens*, el hombre con «conocimiento»—, esta idea debe haber parecido siempre evidente en sí misma. Nadie vio a una lanza fabricar a un cazador a partir de materias primas. Una canción infantil dice «hace falta alguien para conocer algo», pero incluso parece ser un eslogan más persuasivo el que dice «hace falta una cosa más grande para hacer una más pequeña». Sin embargo, un punto de vista inspirado en este eslogan se enfrenta inmediatamente con una cuestión embarazosa, como había señalado Hume: si Dios ha creado y diseñado todas estas cosas maravillosas, ¿quién ha creado a Dios? ¿Superdios? ¿Quién ha creado a Superdios? ¿Supersuperdios? ¿Se ha creado Dios a sí mismo? ¿Es éste un trabajo duro? ¿Le tomó mucho tiempo? ¡No preguntéis! Bien; entonces, puedo preguntar en vuestro lugar si esta blanda aceptación del misterio supone una mejoría sobre la simple denegación del principio de que la inteligencia (o diseño) debe surgir de la Inteligencia. Darwin presentó una explicación que realmente honra la perspicacia de Paley: el trabajo real ha sido el diseño de este reloj y el trabajo no es gratis.

¿Qué cantidad de diseño exhibe una cosa? Nadie ha presentado todavía un sistema de cuantificación del diseño que colme todas nuestras necesidades. El trabajo teórico que se ocupa de esta interesante cuestión está en marcha en varias disciplinas[16], y en el capítulo 6 consideraremos un sistema de medida

natural que facilita una solución inteligente para casos especiales, aunque, mientras tanto, disponemos de un sentido intuitivo y poderoso acerca de las diferentes cantidades de diseño. Los automóviles contienen más diseño que las bicicletas, los tiburones que las amebas e incluso un poema corto contiene más diseño que una frase que dice «Prohibido pisar el césped». (Puedo oír a un lector escéptico que dice «¡Vaya! ¡Más despacio! ¿Es esto algo supuestamente indiscutible?». No, ni de largo. A su debido tiempo intentaré justificar estas afirmaciones pero de momento quiero llamar la atención y basarme en algunas intuiciones familiares, aunque no se consideren dignas de confianza).

La ley de patentes, incluida la ley del *copyright*, es la depositaría de nuestra comprensión práctica de la cuestión. Para justificar la concesión de una patente, ¿cuál sería la cantidad exigida en la innovación del diseño?, ¿cuál sería la cantidad del producto intelectual de otros que uno puede tomar prestada, sin recompensa o reconocimiento? Estas son laderas resbaladizas en las que hemos tenido que construir terrazas arbitrarias codificando lo que, de otra manera, sería materia de interminables disputas. El peso de la prueba en las disputas sobre la cantidad de diseño que es excesiva para ser una mera coincidencia es valorado por nuestra intuición. Nuestras intuiciones son, en este punto, muy fuertes y, como prometo demostrar, razonables. Supongamos que un autor es acusado de plagio y la evidencia es, según se dice, un simple párrafo casi idéntico a otro párrafo de la supuesta fuente: ¿puede tratarse de una simple coincidencia? Depende fundamentalmente de hasta qué punto el párrafo es un formulismo o un lugar común, aunque la mayoría de los pasajes del texto con la longitud de un párrafo son lo bastante «especiales» (como analizaremos enseguida) para determinar la improbabilidad de una creación independiente. En un caso de plagio, ningún jurado razonable exigiría al fiscal que demostrase con exactitud la vía causal a través de la que se produjo la copia alegada. Sobre el acusado recaería, evidentemente, el peso de probar, de manera incuestionable, que el trabajo es un trabajo independiente y no se trata de una copia de algo ya realizado.

El peso de la prueba que le corresponde a un acusado de un caso de espionaje industrial es parecido: ¿es un caso de inocente convergencia de diseños el interior de una nueva línea de artefactos que parece sospechosamente similar en su diseño al de la línea de artefactos del querellante? En tal caso, la única vía para probar realmente la inocencia es presentar clara evidencia de haber realizado el trabajo de I+D necesario (viejos borradores, primeros diseños, primeros modelos y maquetas,

memorias acerca de los problemas encontrados en su realización, etc.). Si faltan tales pruebas, incluso en ausencia de cualquier evidencia física de las actividades de espionaje, el acusado será convicto y merece serlo. No suceden coincidencias cósmicas a tal escala.

El mismo peso de la prueba reina ahora en la biología gracias a Darwin. Lo que llamo el principio de la acumulación de diseños no requiere lógicamente que todos los diseños de este planeta desciendan a través de una rama u otra de un único tronco (raíz o semilla), pero asegura que, dado que cada nuevo objeto diseñado que aparece se debe a una larga inversión en su etiología en algún lugar, la hipótesis más económica será siempre que el diseño ha sido extensamente copiado de otros diseños anteriores, los cuales lo fueron de otros diseños anteriores y así sucesivamente, de tal modo que la innovación real en I+D es minimizada. Sabemos que, de hecho, muchos diseños han sido reinventados independientemente muchas veces —los ojos, por ejemplo, docenas de veces—, aunque todo caso de evolución convergente debe ser probado contra un trasfondo en el que la mayoría de los diseños han sido copiados. Es probable que todas las formas de vida en América del Sur se crearan independientemente de las formas de vida del resto del mundo, pero ésta es una hipótesis muy extravagante que necesitaría ser demostrada, pieza por pieza. Supongamos que descubrimos una nueva especie de ave en alguna isla remota. Aunque no tuviéramos *todavía* evidencia confirmatoria directa de que esta ave está relacionada con las otras aves del mundo, nuestra conjetura está firmemente asegurada, por omisión, desde Darwin, ya que las aves son diseños muy especiales[17]. Así que el hecho de que los organismos —los ordenadores, los libros y otros artefactos— sean efectos de cadenas muy especiales de causas no es, según Darwin, una mera generalización fiable, sino un profundo suceso con el que construir una teoría. Hume reconoció la cuestión —«si se lanzan amontonadas varias piezas de acero al azar nunca se dispondrán de tal modo que puedan componer un reloj»—; así que él y otros pensadores anteriores dedujeron que debían fundamentar este hecho profundo en una Mente. Darwin intuyó cómo distribuirlo en amplios espacios desprovistos de esa invocada Mente, gracias a su idea de que la innovación de diseños puede ser conservada y reproducida, y, por tanto, acumulada.

La idea de que el diseño es algo que ha costado mucho trabajo crear y, por lo tanto, tiene al menos el sentido de que es algo que puede ser conservado (y, por eso, sustraído o vendido) encuentra firme expresión en términos económicos. Si Darwin no hubiera tenido la suerte de haber nacido en un

mundo mercantil que ya había creado a Adam Smith y a Thomas Malthus, no hubiese estado en situación de encontrar piezas prefabricadas que pudiera conjuntar en un nuevo producto con valor añadido. (Como vemos, la idea se aplica muy bien a sí misma). Las diferentes fuentes del diseño que aparecen con la gran idea de Darwin, nos aportan interesantes perspectivas dentro de la propia idea, pero no para disminuir su valor o amenazar su objetividad, del mismo modo que los humildes orígenes del metano no disminuyen sus unidades térmicas británicas (BTU) cuando se utiliza como combustible.

4. *Herramientas para la I+D: ¿«ganchos celestes» o «grúas»?*

El trabajo de I+D no es como palear carbón: es una especie de trabajo «intelectual», y este hecho es el fundamento de otra familia de metáforas que, al mismo tiempo, han seducido y disgustado, iluminado y confundido a los pensadores que se han enfrentado con la «extraña inversión del razonamiento» de Darwin: es decir, la aparente atribución de la inteligencia al mismo proceso de selección natural que Darwin consideraba que no era inteligente.

De hecho, ¿no fue poco afortunado que Darwin decidiera llamar a su principio «*selección* natural», con las connotaciones antropomórficas que supondría esta denominación? ¿No hubiese sido mejor, como Asa Gray le sugirió, reemplazar la imaginería acerca de la «mano que guía la naturaleza» por una discusión sobre las diferentes vías para ganar la carrera de la vida (Desmond y Moore 1991:458)? Mucha gente no se dio cuenta, y Darwin se inclinó a culparse a sí mismo: «Debo de explicar muy mal las cosas», dijo, al tiempo que confesaba: «Supongo que selección natural no es un buen término» (Desmond y Moore 1991:492). Es cierto que este término con dos caras como el dios Jano ha alentado más de un siglo de ardorosos argumentos. Un reciente oponente de Darwin lo resume así:

> La vida en la Tierra, inicialmente entendida como constituyendo *prima facie* una suerte de caso apropiado para un creador, era entrevista, como resultado de la idea de Darwin, como el resultado de un proceso que era, de acuerdo con Dobzhansky, «ciego, mecánico, automático, impersonal» y, según Beer, «derrochador, ciego y desatinado». Pero tan pronto como estas críticas [*sic*] se levantaron contra la selección natural, el «proceso ciego» fue comparado a un poeta, a un compositor, a un escultor o a Shakespeare, a la misma noción de la creatividad que la idea de la selección natural había reemplazado originalmente. Yo creo que hay algo muy, pero que muy erróneo, en tal idea (Bethell 1976).

O algo muy, pero que muy correcto. A los escépticos como Bethell, les parece que hay algo voluntariamente paradójico en denominar al proceso de la

evolución «el relojero ciego» (Dawkins 1986a), porque de esta forma aparta con la mano izquierda («ciega») el discernimiento, el propósito y la providencia que da con la mano derecha. Otros creen que esta manera de hablar —que no es sólo ubicua sino irreemplazable en la biología contemporánea— es precisamente el método correcto para expresar los miles y miles de detallados descubrimientos que la teoría de Darwin ayuda a exponer. No se trata de denegar simplemente la impresionante brillantez de los diseños que se encuentran en la naturaleza. Los biólogos que de vez en cuando se desconciertan con algunos diseños de la naturaleza en apariencia fútiles o torpes, han llegado a comprender eventualmente que han subestimado la ingeniosidad, la diáfana brillantez y la profundidad de la perspectiva que se descubre en una de las creaciones de la madre naturaleza. Francis Crick bautizó maliciosamente esta tendencia con el nombre de su colega Leslie Orgel, citando de lo que él llama «La segunda regla de Orgel: la evolución es más inteligente que tú». (Una fórmula alternativa: ¡La evolución es más inteligente que Leslie Orgel!)

Darwin nos enseñó el modo de llegar desde la «absoluta ignorancia» (como decía su furioso crítico) al genio creativo sin dar por sentada ninguna cuestión, aunque debemos caminar, como veremos, con pies de plomo. Entre las controversias que se arremolinan a nuestro alrededor, la mayoría, si no todas, son diferentes retos a la afirmación de Darwin de que él puede guiarnos a lo largo de todo el camino desde *aquí* (el admirable mundo en el que habitamos) hasta *allí* (el mundo del caos o de la absoluta ausencia de diseño) en el tiempo necesario, sin invocar nada más allá de los mecanismos no inteligentes de los procesos algorítmicos que él había propuesto. Dado que hemos reservado la dimensión vertical de la tradicional pirámide cósmica como una medida (intuitiva) de lo diseñado, podemos escenificar el reto con la ayuda de otra fantasía extraída del folclore.

Gancho celeste. Aeronáutica. Un invento imaginario para la fijación en el cielo; un medio imaginario de suspensión en el cielo (Oxford English Dictionary).

El primer uso de este término, según el *OED*, data de 1915: «Un piloto de avión al que se le ordenó permanecer en su lugar en el aire durante una hora más replicó que “la máquina no dispone de ganchos para colgarse del cielo”». El concepto de «gancho celeste» (*skyhook*) deriva quizá del *deus ex machina* de los antiguos dramaturgos griegos: cuando los escritores de segunda fila se encontraban con que la trama de su obra conducía a los héroes a dificultades

insalvables, solían hacer descender a un dios a escena, como un Superman, para salvar la situación de manera sobrenatural.

Un gancho celeste puede ser también una creación del todo independiente de una evolución folclórica convergente. Sería magnífico disponer de ganchos celestes lo suficientemente grandes como para alzar objetos pesados en circunstancias difíciles y acelerar todo tipo de provectos de construcción. Es triste decirlo, pero son imposibles[18].

Sin embargo, disponemos de grúas. Las grúas pueden hacer el trabajo de elevación que harían nuestros imaginarios ganchos celestes y lo harían de una manera honesta, sin hacerse de rogar. Sin embargo, son costosas. Deben ser diseñadas y construidas con las piezas disponibles y fijadas sobre una base firme en el terreno escogido. Los ganchos celestes son elevadores milagrosos, no soportados e insoportables. Las grúas no son menos excelentes que los elevadores y tienen la ventaja de ser reales. Cualquiera que sea, como yo, un observador de las construcciones en marcha habrá notado con alguna satisfacción que a veces se utiliza una grúa pequeña para emplazar una grúa grande. Y debe haber ocurrido a muchos observadores que, en principio, esta grúa grande puede ser utilizada para poner en situación o acelerar el montaje de una grúa aún más grande. La utilización de grúas en batería es una táctica que raras veces se emplea en los proyectos de construcción del mundo real, pero en principio no hay límites para el número de grúas que pueden disponerse en serie para cumplir algún objetivo.

Ahora imaginemos todas las «elevaciones» que han tenido que ser efectuadas en el espacio de diseño para crear los magníficos organismos y (otros) artefactos que hay en nuestro mundo. Largas distancias han tenido que recorrerse desde los albores de la vida con los entes autorreplicantes más precoces, extendiéndose hacia fuera (diversidad) y hacia arriba (excelencia). Darwin nos había ofrecido un recuento de los más crudos, rudimentarios y estúpidos procesos de elevación, en el borde de la selección natural. Mediante cortos pasos —lo más cortos posibles— este proceso puede atravesar largas distancias gradualmente, a lo largo de los eones. Al menos, así lo dice Darwin. En ningún punto sería necesario algo milagroso desde lo alto. Cada paso ha sido dado mediante una elevación bruta, mecánica, algorítmica, desde la base ya construida con los esfuerzos de las primeras ascensiones.

Este proceso puede parecer increíble. ¿Podría haber sucedido realmente? O bien, ¿necesita el proceso una «ayuda para trepar» ahora y entonces (quizá sólo en el inicio) mediante una especie de gancho celeste o algo parecido? Durante un siglo, los escépticos han tratado de encontrar la prueba que

demuestre que la idea de Darwin no funciona, al menos, *en todo el proceso*. Han estado esperando, buscando afanosamente, rezando por ganchos celestes como excepciones para lo que consideran la sombría visión del algoritmo de Darwin en pleno funcionamiento. De vez en cuando han presentado retos verdaderamente interesantes; saltos e hiatos y otras maravillas que parecen, a primera vista, necesitar ganchos celestes. Pero entonces llegan las grúas descubiertas en muchos casos por los mismos escépticos que esperaban encontrar un «gancho celeste».

Ha llegado el momento de proceder a algunas definiciones más cuidadosas. Un gancho celeste es la fuerza, el poder o el proceso de una «primera mente», una excepción al principio de que todo diseño, o aparentemente diseño, es, en último término, el resultado de una actividad mecánica no inteligente y sin motivo. Una *grúa*, por el contrario, es un subproceso o una característica especial de un proceso de diseño que puede demostrarse que permite la aceleración local del proceso básico y lento de la selección natural, y que igualmente puede demostrarse que es, por sí mismo, el producto predecible (o retrospectivamente explicable) del proceso básico. Algunas grúas son obvias y no son controvertidas; otras son aún discutidas, de manera muy fructífera. Precisamente para darle un sentido general a la amplitud y aplicación del concepto de grúa, me permito exponer tres ejemplos diferentes.

Hoy en día se admite, en general, entre los teóricos de la evolución, que el *sexo* es una grúa. Es decir, las especies que se reproducen sexualmente pueden moverse a través del espacio de diseño con mucha mayor velocidad que la que puede ser conseguida por los organismos que se reproducen asexualmente. Por otra parte, estas especies pueden «discernir» mejoras en el diseño a lo largo del camino de la evolución, las cuales son «invisibles» a los organismos que se reproducen de manera asexuada (Holland 1975). Sin embargo, ésta no puede ser la *raison d’être* del sexo. La evolución no puede ser considerada como un recorrido del camino cuesta abajo, por lo que todo lo que es construido debe tener una inmediata utilidad para compensar el coste. Como han insistido recientemente los teóricos de la evolución, la «elección» de reproducirse sexualmente conlleva un elevado coste *inmediato*: los organismos utilizan solamente el 50 por ciento de sus genes en cualquier transacción (por no mencionar los esfuerzos y riesgos que requiere asegurar una transacción sexual). De este modo, la utilidad a *largo plazo* del incremento de la eficiencia, la precisión y la velocidad del proceso de rediseño —los hechos que hacen que el sexo actúe como una magnífica grúa

—es casi nada para las miópicas competiciones locales que deben determinar qué organismos serán favorecidos en la siguiente generación. Algunos otros beneficios a corto plazo deben haber mantenido la presión de la selección positiva necesaria para hacer de la reproducción sexual una oferta que pocas especies pueden rehusar. Hay una variedad de apremiantes y competitivas hipótesis que pueden resolver este rompecabezas, el cual ha sido enérgicamente planteado a los biólogos por John Maynard Smith [1978]. Para una lucida introducción al estado actual de la cuestión, véase Matt Ridley [1993].

La conclusión que sacamos del ejemplo del sexo es que puede existir una grúa de gran potencia que no se haya creado *con el fin de explotar* este poder, sino por otras razones, aunque su potencia como grúa pueda ayudarnos a explicar por qué se mantiene desde entonces. Una grúa que fue obviamente creada para ser una grúa es la *ingeniería genética*. Los ingenieros genéticos —seres humanos que se ocupan de remendar el ADN recombinante— pueden ahora, sin duda, avanzar a través del espacio de diseño creando organismos que nunca habrían evolucionado por medios «ordinarios».

No se trata de un milagro *si se da por supuesto que los ingenieros genéticos (y los artefactos que utilizan en sus trabajos) son en sí mismos los productos de procesos evolutivos mucho más precoces y lentos*. Si los creacionistas estuvieran en lo cierto, esta humanidad es una especie en sí misma, divina e inaccesible por la vía de Darwin, y entonces la ingeniería genética no sería una grúa después de todo, habiendo sido creada con la ayuda de un gran gancho celeste. No es imaginable que cualquier ingeniero genético piense de sí mismo así, pero es lógicamente una percha disponible, aunque precaria. Menos estúpida, obviamente, es esta idea: si los cuerpos de los ingenieros genéticos son productos de la evolución, pero sus *mentes* pueden hacer cosas creativas que son irreductiblemente no algorítmicas o inaccesibles para todas las vías algorítmicas, entonces los avances de la ingeniería genética pueden implicar la existencia de un gancho celeste. Explorar esta posibilidad será el tema central del capítulo 15.

Una grúa con una historia particularmente interesante es el efecto Baldwin, denominado así por uno de sus descubridores, James Mark Baldwin (1896), aunque fue descubierta más o menos simultáneamente por dos darwinianos de los primeros tiempos, Conwy Lloyd Morgan (conocido por establecer el famoso «canon de parsimonia de Lloyd Morgan») y H. K Osborn (véase Dennett 1983). Baldwin fue un entusiasta darwiniano, abrumado por la perspectiva de que la teoría de Darwin dejara a la mente con

un papel poco importante en el rediseño de organismos. De este modo, preparó el escenario para demostrar que los animales, *en virtud de sus propias actividades inteligentes en el mundo,* podían acelerar o guiar la posterior evolución de sus especies. He aquí lo que se preguntó a sí mismo: ¿cómo puede ser que los animales individuales, resolviendo problemas durante toda su vida, puedan cambiar las condiciones de competición para sus propios descendientes, haciendo que estos problemas sean más fáciles de resolver en el futuro? Y llegó a darse cuenta de que esto era realmente posible, bajo ciertas condiciones; esta conclusión la podemos ilustrar con un simple ejemplo (extraído, con revisiones, de Dennett 1991a).

Supongamos la población de una especie en la que, en el momento de nacer, hay una considerable variación respecto al modelo según el cual sus cerebros tienen dispuestas sus conexiones. Uno de estos modelos de conexión dota a su poseedor de una «buena maña», un talento en su conducta que le protege o incrementa sus oportunidades extraordinariamente. El método estándar para representar tales diferencias de las aptitudes entre los individuos de una población se conoce como «paisaje adaptativo» o «paisaje de aptitudes» (S. Wright 1931). La altitud en ese diagrama señala la idoneidad (mientras más alto mejor), y la longitud y la latitud representan algunos factores del diseño individual (en este caso, hechos relacionados con el «cableado cerebral»). Cada modelo diferente según el cual un cerebro puede ser «cableado» se representa por una de las varillas que componen el paisaje: cada varilla es un *genotipo* diferente. El hecho de que precisamente una de las combinaciones es la buena —esto es, mejor que la media— se ilustra por el modo en que aparece como un poste telefónico en un desierto.

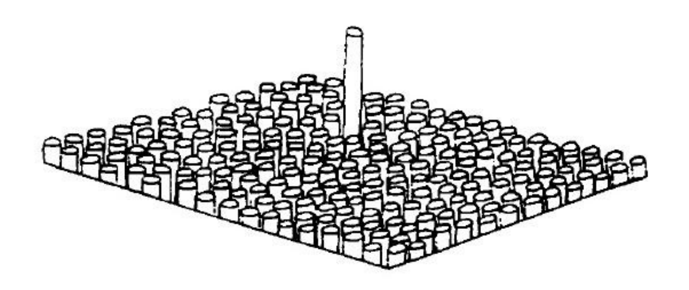


Figura 3.1

Como la figura 3.1 demuestra claramente, tan sólo se favorece un «cableado»; los otros, no importa lo «cerca» que estén de ser un buen cableado, son iguales en cuanto a aptitudes. Naturalmente, este pico aislado es como una

aguja en un pajar: que sería prácticamente invisible a la selección natural. Aquellos escasos individuos en la población que son lo bastante afortunados para tener el genotipo de la «buena maña» tendrían dificultades para hacerlo pasar a sus descendientes, dado que en la mayoría de las circunstancias, sus oportunidades de encontrar una pareja que también tenga el genotipo de la «buena maña» son remotas, con un error de tiro de casi una milla.

Ahora vamos a introducir un cambio «menor»: supongamos que aunque los organismos individuales *comiencen* con diferentes «cableados» (cualquier cableado fue ordenado por su particular genotipo o receta genética) —como se muestra por su dispersión en el paisaje de aptitudes—, tienen alguna capacidad para ajustar o revisar su «cableado» dependiendo de lo que encuentren durante su vida. (En el lenguaje de la teoría de la evolución, se dice que existe alguna «plasticidad» en su *fenotipo*. El fenotipo es el diseño eventual del cuerpo creado por el genotipo en interacción con el medio ambiente. Gemelos idénticos criados en diferente medio ambiente comparten un genotipo que puede ser extraordinariamente diferente en el fenotipo). Supongamos, pues, que estos organismos pueden terminar, después de la exploración de su entorno, con un diseño diferente del que tenían al nacer. Podemos suponer que sus exploraciones son aleatorias, aunque poseen una capacidad innata para reconocer (y retener) una «buena maña» dada cuando tropiezan con ella. De este modo, aquellos individuos que comienzan su vida con un genotipo que es parecido al genotipo de la «buena maña» —se diferencian por detalles mínimos en su diseño— es más probable que terminen por coincidir en el diseño que aquellos otros que han nacido con un diseño muy alejado.

Este comienzo en la cabeza de la carrera para rediseñarse a sí mismos puede darles una pequeña ventaja cuando se produzca la «quiebra» malthusiana, si la «buena maña» es tan buena que aquellos que nunca la han aprendido o lo hicieron «demasiado tarde» se encuentran en grave desventaja. En poblaciones con este tipo de plasticidad fenotípica, un pequeño error de tiro es *mejor* que cuando el error es de una milla. Para esta población, el poste de teléfono en el desierto llega a ser la cima de una colina que asciende gradualmente, como se representa en la figura 3.2; aquellos situados cerca de la cima, aunque comienzan con un diseño que les sirve no mejor que otros, tenderán a descubrir el diseño de la cima en poco tiempo.

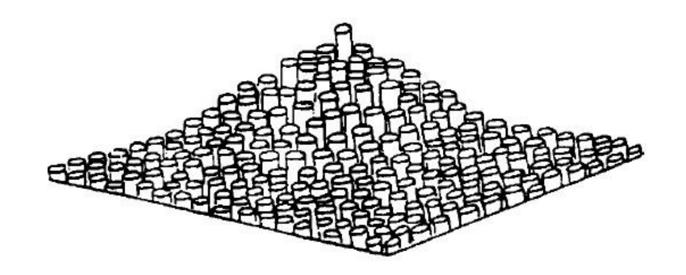


Figura 3.2

A largo plazo, la selección natural —rediseño a nivel del genotipo— *tenderá a seguir esta línea y a confirmar* las direcciones tomadas por las sucesivas exploraciones llevadas a cabo por los organismos individuales: rediseño a nivel individual o fenotipo.

El modo según el cual he descrito el efecto Baldwin conserva ciertamente la mente hasta un grado mínimo, si no la elimina del escenario; todo lo que esto requiere es una capacidad bruta, mecánica, para detener un paseo aleatorio en cuanto aparece la «buena maña», una mínima capacidad para «reconocer» un pequeñísimo progreso, para «aprender» algo mediante el método ciego de la prueba y el error. En efecto, he descrito este asunto en términos *conductistas*. Lo que Baldwin descubrió era que las criaturas capaces de realizar un «aprendizaje reforzado», no sólo actúan mejor individualmente que las criaturas que están enteramente cableadas: sus especies evolucionan más rápidamente por su mayor capacidad para descubrir mejoras en los diseños en la vecindad**[19]**. Pero no era así como Baldwin describió el efecto que proponía. Su temperamento estaba muy alejado del conductismo. Como Richards señala:

El mecanismo se conforma con las hipótesis ultradarwinianas; no obstante, permite a la conciencia y a la inteligencia un papel en la dirección de la evolución. Por disposición filosófica y por convicción, Baldwin era un metafísico espiritualista. Sentía el latido de la conciencia en el universo, que latía a lo largo de todos los niveles de la vida orgánica. Baldwin comprendió el poder de las explicaciones mecanicistas de la evolución (R. J. Richards 1987:480)**[20]**.

El efecto Baldwin, bajo diferentes nombres, ha sido descrito de diversas formas, defendido y censurado a lo largo de los años, y recientemente redescubierto de modo independiente varias veces más (por ejemplo, Hinton y Nowlan 1987).

Aunque ha sido descrito y reconocido en los textos de biología con regularidad, también ha sido habitualmente evitado por los pensadores

demasiado cautelosos, que han creído que tiene un regusto de la herejía lamarckiana (la presumida posibilidad de herencia de las características adquiridas; véase el capítulo 11 para una discusión detallada). Este rechazo es especialmente irónico dado que, como señala Richards, Baldwin intentó conseguir un *sustituto* aceptable —y realmente lo es— de los mecanismos lamarckianos.

El principio ciertamente pareció acabar con el lamarckismo, al mismo tiempo que aportaba este factor positivo en la evolución que incluso los fieles darwinistas como Lloyd Morgan anhelaban. Y para aquellos con apetito metafísico, revelaba que bajo la resonante y mecánica cobertura de la naturaleza darwiniana, podía encontrarse la mente (R. J. Richards 1987:487).

En definitiva, no la mente —si con ello entendemos la mente como un gancho celeste totalmente desarrollado, intrínseco y original—, sino una mente excelente, mecanicista y conductista, que actúa como una grúa. Pero esto no significa nada; Baldwin descubrió un efecto que aumenta el poder —localmente— del subyacente proceso de selección natural allí donde éste opere. Muestra cómo el proceso «ciego» de este fenómeno básico de la selección natural puede ser inducido por una limitada cantidad de «impulso hacia delante» en las actividades de los organismos individuales, las cuales crean diferencias de aptitudes sobre las que la selección natural puede entonces actuar. Esta es una complicación bienvenida, una arruga en el tejido de la teoría de la evolución que elimina una razonable y exigente fuente de duda, y potencia nuestra visión acerca del poder de la idea de Darwin, especialmente cuando ésta es vertida en las múltiples aplicaciones que de ella se han derivado. Esto es característico del resultado de otras investigaciones y controversias que vamos a analizar: la motivación, la pasión que dirigió la investigación era la esperanza de encontrar un gancho celeste; el triunfo fue hallar que el mismo trabajo podía ser hecho con una grúa.

5. *¿Quién teme al reduccionismo?*

> Reduccionismo es una palabra indecente y una especie de farisaico «más papista que el Papa» se ha puesto de moda.
>
> Richard Dawkins, *The Extended Phenotype*

En este tipo de conflicto, el término «reduccionismo» es el que más a menudo se encuentra en boca de todos, como una expresión típicamente injuriosa. Los que sueñan con la existencia de ganchos celestes llaman «reduccionistas» a

los que se decantan con vehemencia por las grúas y hacen que el reduccionismo aparezca con frecuencia como un filisteo y empedernido, o dicho llanamente, como el mismísimo demonio. Pero como la mayoría de los términos injuriosos, la palabra «reduccionismo» no tiene un significado preciso. La imagen central es la de alguien que sostiene que una ciencia se «reduce» a otra: que la química se reduce a la física, que la biología se reduce a la química, que las ciencias sociales se reducen a la biología, por ejemplo. El problema es que existen, al mismo tiempo, lecturas prudentes y lecturas disparatadas del reduccionismo. Según las lecturas prudentes, es posible (y deseable) *unificar* la química y la física, la biología y la química e, incluso, las ciencias sociales y la biología. Después de todo, las sociedades están formadas por seres humanos, que, como mamíferos que son, entran en los principios de la biología que se extienden a todos los mamíferos. Los mamíferos, a su vez, están formados por moléculas que obedecen a las leyes de la química y ésta, a su vez, a las reglas de la física subyacente. Ningún científico en su sano juicio discutiría hoy esta lectura prudente; reunidos en asamblea, los jueces de la Corte Suprema están vinculados a la ley de la gravedad como cualquier avalancha, ya que son, en último término, un conjunto de objetos físicos. Según las lecturas disparatadas, los reduccionistas deben abandonar los principios, teorías, vocabularios y leyes científicas de alto nivel y sustituirlos por términos de nivel más bajo. Según esta lectura descabellada, un sueño reduccionista podría ser escribir «una comparación entre Keats y Shelley desde el punto de vista molecular», o «el papel de los átomos de oxígeno en la economía de los suministros», o bien una «explicación de las decisiones de la corte de Rehnquist según las fluctuaciones de la entropía». Probablemente nadie es un reduccionista en este sentido disparatado y todo el mundo lo es en su versión prudente, de modo que la «acusación» de reduccionismo es demasiado vaga para merecer una respuesta. Si alguien nos dice «¡pero eso que dice es muy reduccionista!», haríamos bien en responderle: «¡Esa es una queja pintoresca pasada de moda. ¿Qué piensa de la Tierra?».

Me satisface poder decir que en los últimos años, algunos de los pensadores que más admiro han salido en defensa de una u otra versión del reduccionismo, cuidadosamente restringida. El científico cognitivo Douglas Hofstadter, en Gödel, Escher, Bach, compuso un «Preludio… fuga de la hormiga» (Hofstadter 1979:273-336) que es un himno analítico a las virtudes del reduccionismo en su lugar apropiado. George C. Williams, uno de los evolucionistas preeminentes de nuestros días, publicó *Una defensa del*

reduccionismo en la biología evolucionista (1985). El zoólogo Richard Dawkins ha distinguido lo que él llama el reduccionismo jerárquico o gradual del reduccionismo del precipicio; rechaza solamente la versión del precipicio (Dawkins 1986b:74).[21] Más recientemente el físico Steven Weinberg, en *El sueño de una teoría final* (1992), ha escrito un capítulo titulado «Dos brindis por el reduccionismo» en el que distingue el reduccionismo no comprometido (algo malo) y el reduccionismo comprometido (al que apoya encarecidamente). Esta es mi propia versión. Hay que distinguir entre el reduccionismo que, en general, es algo bueno, del *reduccionismo voraz*, que no lo es. En el contexto de la teoría de Darwin, la diferencia es simple: los reduccionistas voraces piensan que todo puede explicarse sin grúas; los buenos reduccionistas creen que todo puede explicarse sin ganchos celestes.

No hay razón para ser transigente con lo que yo llamo el buen reduccionismo. Es simplemente el compromiso con la ciencia que no da por sentado lo que no ha sido probado, sin dejarse engañar al principio aceptando misterios o milagros. (Para otra perspectiva sobre este tema véase Dennett 1991a:33-39). Brindemos por esta rama del reduccionismo; estoy seguro de que Weinberg estaría de acuerdo. Pero en sus ansias por conseguir un pacto, en su celo por explicar demasiadas cosas y con demasiada rapidez, a menudo los científicos y los filósofos subestiman las complejidades, tratando de saltarse capas completas o niveles de la teoría en su apresurada carrera por anclar todas las cosas, con seguridad y claridad, a sus fundamentos. Este es el pecado del reduccionismo voraz, aunque hay que señalar que este reduccionismo debe ser condenado sólo cuando la ansiedad y la voracidad conducen a la falsificación de los fenómenos. En sí mismo, el deseo de reducir y unificar para explicar todo en una gran teoría que acoja la totalidad, no tiene más motivo para ser condenado como inmoral que el impulso contrario que condujo a Baldwin a su descubrimiento. No es erróneo apostar por teorías simples o por fenómenos que ninguna teoría simple (¡o compleja!) haya podido explicar; lo que es erró neo es una interpretación entusiasta y falsa en una u otra dirección.

La peligrosa idea de Darwin es reduccionismo encarnado[22], con la promesa de unir y explicar todas las cosas en una visión espléndida. Esta idea, al tratarse de un proceso algorítmico, se hace más poderosa, dado que la neutralidad del sustrato que posee nos permite considerar sus aplicaciones a cualquier cosa. No respeta los límites materiales. Se aplica, como ya hemos empezado a ver, incluso a sí misma. El temor más frecuente respecto a la idea de Darwin no es que justamente no explique, sino que *disculpe dando*

explicaciones a las mentes los propósitos y los significados que todos nosotros hemos estimado. La gente teme que una vez que el ácido universal ha penetrado en los monumentos que admiramos, éstos dejarán de existir, disueltos en un irreconocible y nada atractivo fango de destrucción científica. Esto no es un temor bien fundamentado; una *apropiada* explicación reduccionista de estos fenómenos los dejaría aún en pie, aunque desmitificados, unificados y asentados en cimientos más seguros. Podemos aprender algunas cosas sorprendentes e incluso chocantes acerca de estos tesoros, a menos que la evaluación que hagamos de estas cosas estuviera basada en confusión o errónea identidad. ¿Cómo puede el incremento de nuestro conocimiento disminuir su valor a nuestros ojos?**[23]**

Un temor más razonable y realista sería que el disparatado abuso del razonamiento de Darwin pueda conducirnos a negar la existencia de niveles reales, de complejidades reales y de fenómenos reales. Con nuestros propios esfuerzos mal dirigidos podemos llegar a descartar o a destruir algo valioso. Hay que trabajar duro para mantener apartados estos dos temores y comenzar por el reconocimiento de las presiones que tienden a distorsionar la propia descripción de los hechos. Por ejemplo, entre los muchos que se encuentran incómodos con la teoría de la evolución, existe una fuerte tendencia a exagerar la cantidad de desacuerdos entre los científicos («no es más que una teoría y hay muchos científicos famosos que no la aceptan»), al tiempo que no debemos exagerar la contra-argumentación al describir «lo que la ciencia ha demostrado». Por esta vía encontraremos muchos ejemplos de genuinos desacuerdos científicos y cuestiones inestables. No encuentro una razón para ocultar o disminuir la importancia, de estas dificultades, no importa la manera en que hayan aparecido, y una cierta cantidad de trabajo corrosivo ya ha sido realizado por la peligrosa idea de Darwin, y no puede nunca ser arruinado.

Ha llegado la hora de ser capaces de ponernos de acuerdo en un resultado. Incluso si la relativamente modesta idea de Darwin acerca del origen de las especies llegara a ser rechazada por la ciencia —sí, absolutamente desacreditada y sustituida por una visión mucho más poderosa (y por el momento inimaginable)—, aún seguiría siendo una convicción irremediable y debilitada en cualquier defensor ponderado de la tradición expresada por Locke. Esto se ha logrado al abrir nuevas posibilidades de imaginación y al destruir por completo cualquier ilusión que uno podría haber tenido acerca de los fundamentos de un argumento como la prueba *a priori* de Locke sobre el carácter inconcebible del diseño sin mente. Antes de Darwin esto era tan inconcebible, en sentido peyorativo, que nadie se atrevía a tomar la hipótesis

en serio. Probarla ya era otro asunto, pero la evidencia aumenta y es cierto que nosotros podemos y debemos tomarla seriamente. De este modo, sea lo que sea lo que podamos pensar sobre el argumento de Locke, éste es ahora tan obsoleto como la pluma de ave con la que fue escrito, una fascinante pieza de museo, una curiosidad que no puede realmente funcionar en el mundo intelectual de nuestros días.

CAPÍTULO 4
El árbol de la vida

I. *¿Cómo imaginarnos el árbol de la vida?*

> La extinción tan sólo ha separado grupos: de ninguna manera los ha hecho; si toda forma que ha vivido en esta tierra reapareciera súbitamente, aunque sería completamente imposible dar definiciones con las que cada grupo se distinguiera de los otros grupos —ya que todos se mezclarían entre sí por etapas, de forma tan excelente como las de las variedades existentes más excelentes—, a pesar de todo sería posible una clasificación natural o, al menos, una disposición natural.
>
> Charles Darwin,
> *El origen de las especies.*

En el capítulo anterior, la idea del trabajo de Investigación + Desarrollo (I+D) entendido como una exploración realizada dentro de lo que he llamado el espacio de diseño, fue introducida al vuelo, sin prestar atención a los detalles o a las definiciones de los términos utilizados. Con el fin de diseñar el cuadro global, he utilizado varias declaraciones controvertidas con la promesa de que más tarde las defendería. Dado que la idea del espacio de diseño va a ser muy utilizada, ahora es cuando debo asegurarme, siguiendo el discurso de Darwin, de que comenzaré, una vez más, en el medio, examinando, en primer lugar, algunos modelos *reales* en algunos espacios relativamente bien explorados. Estos modelos servirán como guías, en el próximo capítulo, para conseguir una perspectiva más general de los modelos *posibles*, y la vía por la que cierto tipo de procesos tienen posibilidades de hacerse realidad.

Consideremos el árbol de la vida, es decir, el gráfico que representa las trayectorias en el tiempo de todas las cosas que han vivido en este planeta o —en otras palabras— la totalidad del abanico de las *descendencias*. Las reglas para dibujar el árbol de la vida son simples. La línea del tiempo de un organismo comienza cuando nace y termina cuando muere, y además pueden surgir o no surgir de esta línea del tiempo las líneas que corresponden a los descendientes. Una imagen ampliada de las líneas de descendencia de un organismo —si existe alguna— variará de aspecto de acuerdo con varios

factores: si el organismo se reproduce por fisión o por gemación, o mediante huevos o con crías vivas (vivíparo) y si el organismo progenitor sobrevive de modo que coexistirá durante cierto tiempo con su descendencia. Pero estos microdetalles del abanico, por lo general, no nos conciernen ahora. No existe una controversia importante sobre la afirmación de que toda la diversidad de la vida que ha existido en este planeta se deriva de este singular abanico; por el contrario, la controversia surge cuando se plantea cómo descubrir y describir, *en términos generales*, las diferentes fuerzas, principios, limitaciones, etc., que nos permiten dar una explicación científica de los modelos, en toda su diversidad.

La Tierra tiene alrededor de 4.500 millones de años y las primeras formas de vida aparecieron muy «pronto»; los organismos unicelulares simples —los *procariotas*-aparecen al menos hace 3.500 millones de años, y probablemente durante 2.000 millones de años toda la vida estuvo representada por bacterias, algas verdeazuladas y formas similares. Hace 1.400 millones de años se produjo una gran revolución: algunas de estas formas de vida simples unieron sus fuerzas cuando algunas bacterias, como los procariotas, penetraron a través de las membranas de otros procariotas creando los *eucariotas*, células con núcleo y otros corpúsculos internos especializados (Margulis 1981). Estos corpúsculos internos, llamados *organelas* o *plásmidos*, fueron la innovación clave del diseño que abrió las regiones del espacio de diseño habitadas hoy. En las plantas, los *cloroplastos* son responsables de la fotosíntesis y las *mitocondrias*, que se encuentran en todas las células de las plantas, animal u hongo —todo organismo con células nucleadas—, son las factorías fundamentales que procesan el oxígeno para obtener la energía que nos permite mantener a raya la segunda ley de la termodinámica, mediante la explotación de los materiales y de la energía que nos rodea. El prefijo *eu* en griego significa «bueno» y desde nuestro punto de vista, los eucariotas son, ciertamente, una mejora, dado que, gracias a su complejidad interna, pueden especializarse, lo que hizo posible la creación de organismos multicelulares, como nosotros mismos.

La *segunda revolución* —la aparición del primer organismo multicelular — tuvo que esperar 700 millones de años, más o menos. Una vez que los organismos multicelulares aparecieron en escena, el ritmo se aceleró y el subsiguiente abanico de plantas y animales —desde helechos y flores a insectos, reptiles, aves y mamíferos— ha poblado el mundo de hoy con millones de especies diferentes. En el proceso, millones de otras especies hicieron su aparición y han desaparecido. Seguro que se han extinguido

muchas más especies de las que ahora existen, quizá cien especies extinguidas por cada especie existente.

¿Cuál es la forma global de este gran árbol de la vida que ha ido extendiendo sus ramas a lo largo de 3.500 millones de años? ¿Qué aspecto tendría si, con el ojo de Dios, lo pudiéramos contemplar de una vez, extendiéndose ante nosotros continuamente, en una dimensión espacial? En los gráficos científicos, la práctica habitual es situar al tiempo en el eje horizontal, con las formas *iniciales* de la vida a la izquierda y las tardías a la derecha, pero los diagramas de la evolución han sido siempre una excepción, al colocar habitualmente al tiempo en la dimensión vertical. Aún es más curioso que nos hayamos acostumbrado a dos convenciones opuestas, colocando las etiquetas en la dimensión vertical; y con estas convenciones han llegado sus correspondientes metáforas. Podemos ubicar a las formas más *precoces* arriba y a las *tardías* abajo, en cuyo caso nuestros diagramas mostrarán los ancestros y sus *descendientes*. Darwin utilizó esta convención cuando hablaba de especiación como modificación con descenso, y, desde luego, en el título de su obra sobre la evolución humana, *The Descent of Man, and Selection in Relation to Sex* (*El origen del hombre*, 1871). De manera alternativa, podemos dibujar un árbol con su orientación normal en el espacio, de modo que aparezca realmente como un árbol en el que los últimos «descendientes» en el tiempo componen las ramas y las ramillas que se *elevan*, a lo largo del tiempo, desde el tronco y las raíces iniciales. Darwin también utilizó esta convención —por ejemplo, en el único diagrama de su libro sobre *El origen de las especies*—, pero también al usar expresiones que alinean lo *más alto* con lo que aparece *más tarde*. Ambos grupos de metáforas coexisten con escasas dificultades en el lenguaje y en los diagramas de la biología actual. (Esta tolerancia ante una imaginería confusa no se limita a la biología. Los «árboles genealógicos» son, a menudo, dibujados con los antepasados en la parte alta, y los lingüistas generativos, entre otros, dibujan sus árboles derivativos boca abajo con la «raíz» en la parte superior de la página).

Dado que ya había propuesto marcar la dimensión vertical en el espacio de diseño como una medida de la cantidad de diseño, de modo que mientras *más alto* equivale a *más diseñado*, debemos tener cuidado en anotar que en el árbol de la vida (dibujado desde el lado derecho hacia arriba, como he propuesto hacer) *más alto* equivale a *más tarde* (y nada más). No significa *necesariamente* más diseñado. ¿Cuál es la relación entre tiempo y diseño o cual podría ser? ¿Las cosas que han sido más diseñadas pueden aparecer

primero y perder diseño gradualmente? ¿Es posible un mundo en el que las bacterias sean las descendientes de los mamíferos y no al revés? Estas cuestiones relacionadas con posibilidades serían mucho más fáciles de responder si analizamos, en primer lugar y un poco más de cerca, lo que ha sucedido realmente en nuestro planeta. Hay que aclarar que, por lo pronto, la dimensión vertical en el diagrama reproducido abajo se relaciona con el tiempo y sólo el tiempo, con lo más *precoz* en la parte inferior y lo más *tardío* en la parte superior. Siguiendo la práctica estándar la dimensión izquierda-derecha se considera a modo de sumario de la diversidad en un solo plano. Cada organismo individual debe tener su línea del tiempo, distinta de la de los otros organismos, de modo que incluso si dos organismos son exactamente duplicados uno del otro, átomo por átomo, deben aparecer en el árbol lado con lado en el mejor de los casos. No obstante, como están alineados en sentido ascendente, puede hacerse de acuerdo con alguna medida o conjunto de medidas de las diferencias en la forma corporal individual —*morfología*—, para usar el término técnico.

Así, volviendo a nuestra pregunta, ¿cuál sería la forma global del árbol de la vida si pudiéramos verlo con una sola mirada? ¿No podría parecerse a una palmera, como en la figura 4.1?

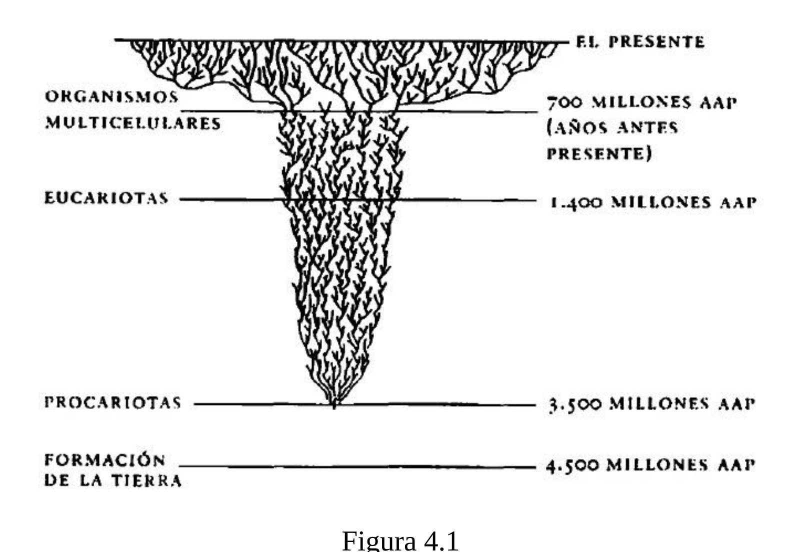


Figura 4.1

Este es el primero de muchos árboles o *dendrogramas* que tendremos que considerar y, naturalmente, la resolución limitada de la tinta en las páginas difuminarán millones y millones de líneas separadas que aparecerán como juntas. Por el momento, he dejado deliberadamente la «raíz» del árbol borrosa e imprecisa. Todavía estamos explorando la zona media del árbol, reservando

los comienzos para un posterior capítulo. Si logramos una imagen ampliada del tronco del árbol y pudiésemos examinar una sección transversal —un «instante» en el tiempo—, veríamos miles y miles de millones de organismos individuales unicelulares, de una fracción de estos partirían líneas que nos conducirían a una progenie situada ligeramente más arriba en el tronco. (En los primeros días de la vida, la reproducción se hacía por gemación o fisión; algo más tarde, evolucionó un tipo de sexo unicelular, aunque la dispersión del polen por el aire y la puesta de huevos y los otros fenómenos relativos a nuestra reproducción sexual deberían esperar a la revolución multicelular que ocurre en la fronda del árbol de la vida). Habría alguna diversidad y alguna revisión de los diseños a lo largo del tiempo, de modo que quizá la totalidad del tronco aparecería inclinándose a derecha o a izquierda o extendiéndose más de lo que hemos mostrado. ¿Es precisamente nuestra ignorancia la que nos impide diferenciar corrientes dominantes en este «tronco» de variedades unicelulares? Quizás aparecería el árbol con varias ramas de suficiente grosor para ser visibles, con su extremo cerrado, como en la figura 4.2, como señal de experimentos en diseños unicelulares alternativos, acaecidos durante varios cientos de millones de años que terminarían eventualmente por extinguirse.

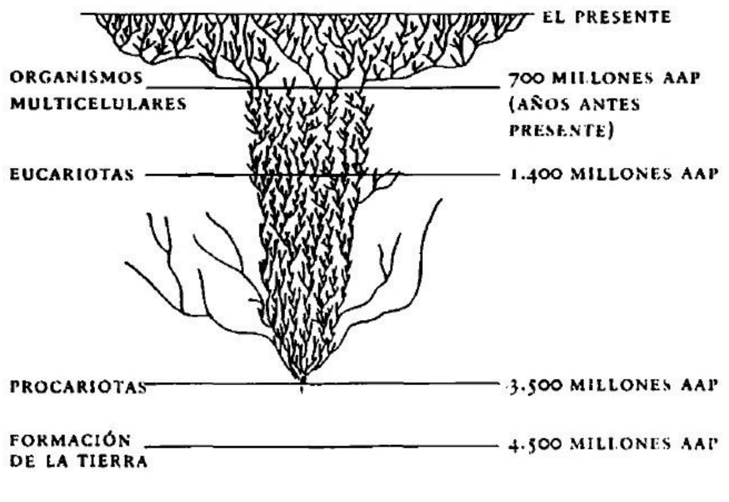


Figura 4.2

Debe haber habido miles de millones de fallidos experimentos de diseños, pero quizá ninguno aún muy alejado de la norma de un simple organismo unicelular. En cualquier caso, si ampliáramos la imagen del tronco, observaríamos un crecimiento abundante de alternativas de vida corta, como en la figura 4.3, pero todas invisibles, contra la norma de la replicación conservadora. ¿Cómo podemos estar seguros de esto? Porque, como veremos,

las probabilidades están totalmente en contra de que cualquier mutación sea más viable que el tema del cual dicha mutación es una variación.

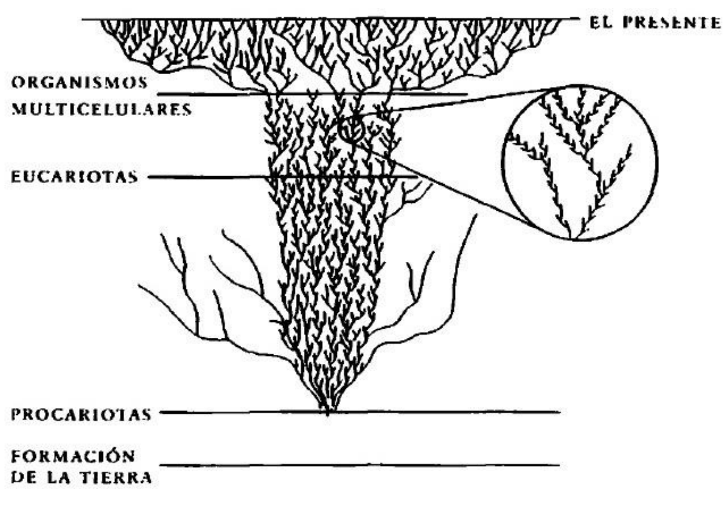


Figura 4.3

Hasta que se inventa la reproducción sexual, casi todas las ramas que observamos, bajo cualquier nivel de ampliación, divergen. No obstante, las excepciones son llamativas. En el tiempo de la revolución eucariota, si observamos precisamente en el lugar adecuado, veremos una bacteria que penetra en el cuerpo rudimentario de algún otro procariota para crear el primer eucariota. Su progenie tendrá una herencia doble; contiene dos secuencias totalmente independientes de ADN, una de la célula huésped y otra de la célula «parásita», que comparte su suerte con la célula huésped, y ligando la suerte de todos sus descendientes (ahora en el camino de llegar a ser una benigna mitocondria residente) a la suerte de las células que habitarán los descendientes de la célula que fue invadida por primera vez. Es un hecho sorprendente de la geometría microscópica del árbol de la vida: linajes totales de mitocondrias, delicadas cosas vivientes por derecho propio, con su propio ADN, viviendo todas sus vidas entre las paredes de las células de organismos más grandes que forman otros linajes. En principio, esta penetración sólo ha sucedido una vez, pero podemos suponer que hayan ocurrido muchas experiencias de este tipo de simbiosis tan radical (Margulis 1981; sumarios accesibles se encontrarán en Margulis y Sagan 1986 y 1987).

Muchos millones de años después, una vez que la reproducción sexual fue establecida, en lo alto del frondoso ramaje de nuestro árbol de la vida (el sexo ha evolucionado aparentemente muchas veces, aunque hay desacuerdo en este tema), si observamos con lente de aumento, y de cerca las trayectorias vitales

de los organismos individuales, encontraremos un tipo diferente de vinculación entre los individuos —apareamientos— que tiene como resultado la aparición de descendientes como estallidos de estrellas. Al incrementar la ampliación de las imágenes y «mirar a través del microscopio» podemos observar en la figura 4.4 cómo, al contrario de lo que sucede en la conjunción que ha creado los eucariotas, donde las dos secuencias del ADN son preservadas en su totalidad y mantenidas distintas dentro de los cuerpos de la progenie, en cada apareamiento sexual cada progenie mantiene su propia y única secuencia de ADN, tejida conjuntamente mediante un proceso que recoge el 50 por ciento de uno de los progenitores y el 50 por ciento del otro. Como es natural, cada célula de la progenie también contiene mitocondrias y éstas siempre provienen de un solo progenitor, la hembra (si es varón, las mitocondrias de sus células se encuentran en un callejón sin salida desde el punto de vista evolutivo; no pasarán a sus descendientes, que conseguirán las mitocondrias de su madre). Dejemos por el momento el examen con ampliación de los apareamientos con progenie y observemos (en la figura 4.4) que la mayoría de las trayectorias vitales de estos descendientes terminan sin apareamiento o, al menos, sin descendencia propia. Esta es la llamada «quiebra» malthusiana. Dondequiera que miremos, las ramas y ramillas están recubiertas de una especie de pelusa terminal formada por breves trayectorias vitales, de nacimiento-muerte, sin más progreso.

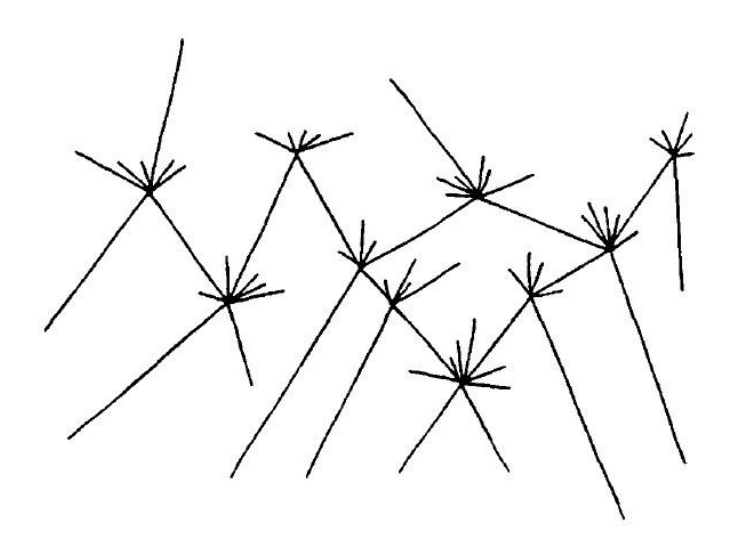


Figura 4.4

Sería imposible observar al mismo tiempo todos los puntos de ramificación y de unión en la totalidad del árbol de la vida a lo largo de 3.500 millones de años, pero si dejamos los detalles y contemplamos el árbol en una amplia

perspectiva, reconoceremos detalles muy familiares. Muy pronto, en el abanico multicelular que comenzó hace aproximadamente unos 700 millones de años a partir del tronco de los organismos unicelulares, podemos ver la horquilla de la divergencia que crea dos grandes ramas —los reinos animal y vegetal— y otra para los hongos, que salen del tronco de los organismos unicelulares. Y si miramos más de cerca, veríamos que, una vez que la separación ha alcanzado cierta distancia, no se encuentra ningún apareamiento que reúna las trayectorias de sus miembros individuales. A esta altura del tiempo evolutivo, los grupos han quedado reproductivamente aislados y el hiato entre estos grupos se hace cada vez más y más amplio[24]. Posteriores horquillas en el árbol de la vida crean los *phyla*, los órdenes, las clases, las familias, los géneros y las especies multicelulares.

2. *Codificando con colores una especie en el árbol de la vida*

¿Qué es lo que hace que una *especie* aparezca como tal en este árbol? Dado que la cuestión de lo que es una especie, o cómo una especie comienza, continúa generando controversias, aprovecharemos la ventaja que tiene la perspectiva del ojo de Dios, que hemos adoptado temporalmente, para observar más de cerca la totalidad del árbol de la vida y comentar lo que sucedería si intentásemos codificar con un color concreto una especie singular. Una cosa puede asegurarse: cualquiera que sea la región que marquemos con un color, será una región individual y conectada. Pequeñas masas no separadas de organismos, no importa lo similares que sean en apariencia y en morfología, pueden considerarse como compuestas de miembros de una especie *singular* que deben unirse por la descendencia. El siguiente punto que conviene subrayar es que hasta que la reproducción sexual entra en escena, el sello distintivo del aislamiento reproductivo no puede ser tenido en cuenta. Esta útil condición que marca los límites no tiene definición en el mundo asexuado. En aquellos filamentos antiguos y contemporáneos del árbol de la vida que se reproducen asexualmente, los agrupamientos de un tipo o de otro pueden interesarnos por varias razones —agrupamientos por morfología o por conducta compartidas o similaridad genética, por ejemplo— y podemos decidir denominar como una especie al grupo resultante, aunque puede que no existan límites teóricos definidos con precisión que delimiten tal especie. Así que vamos a concentrarnos en las especies de reproducción sexual, todas ellas situadas en las frondosidades

multicelulares del árbol de la vida. ¿Cómo podemos conseguir colorear de rojo todas las líneas de la vida de una especie singular? Podemos comenzar examinando aleatoriamente los organismos individuales hasta que encontremos uno con bastantes descendientes. Lo llamaremos Lulu y lo teñiremos de rojo (el rojo está representado con líneas gruesas en la figura 4.5). Ahora iremos coloreando de rojo poco a poco y en dirección ascendente todos los descendientes de Lulu; éstos serán todos miembros de una especie, *a menos* que comprobemos que nuestra tinta roja se extiende a dos ramas más altas que son distintas, ninguno de cuyos miembros establecen conexiones a través del vacío. Si esto sucede, sabremos que se ha producido una especiación y volveremos atrás para tomar varias decisiones. En primer lugar, debemos escoger si vamos a mantener una de las ramas de color rojo (la especie «progenitora» sigue siendo roja y la otra rama se considera la nueva especie hija) o si detenemos la tinción tan pronto como ocurre la divergencia o ramificación (la especie «progenitora» se ha extinguido fisionándose en dos especies hijas).

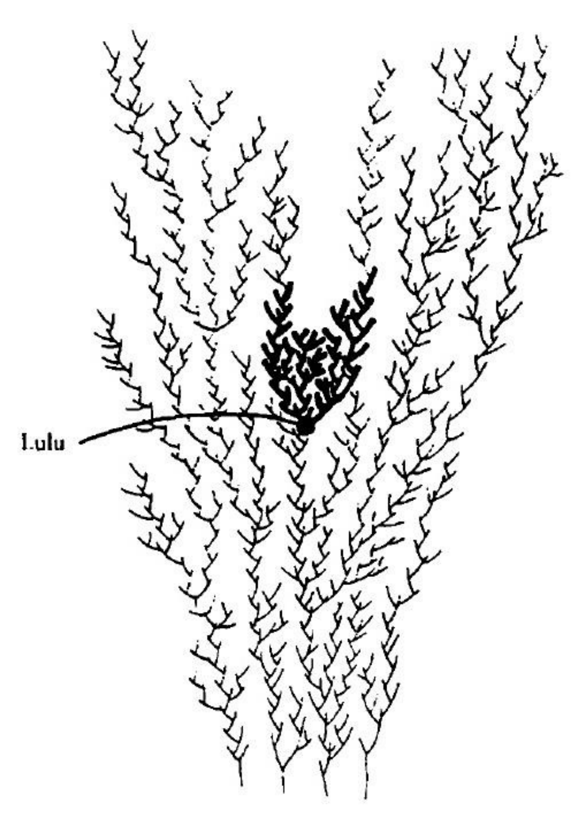


Figura 4.5

Si los organismos en la rama de la izquierda son muy similares, en apariencia, dotación y hábitos, a los contemporáneos de Lulu, mientras que

los organismos de la rama derecha presentan casi todos nuevos cuernos, pies palmeados o rayas, en este caso parece bastante obvio que debemos señalar la rama izquierda como la continuación de la especie progenitora y la rama derecha como el nuevo retoño. Si ambas ramas muestran pronto modificaciones mayores, nuestra decisión, basada en el código rojo, deja de ser tan obvia. No hay hechos secretos que puedan indicarnos qué elección es la correcta, ni qué elección «moldea a la naturaleza en sus vinculaciones», porque estamos observando directamente los lugares donde las vinculaciones tendrían que estar y no están. Para ser una especie basta con ser una de estas ramas de organismos entre las que tienen lugar cruzamientos, y para ser *coespecífico*, esto es, de la misma especie de algún otro organismo (contemporáneo o no) basta con formar parte de la misma rama. La elección que hacemos dependerá entonces de consideraciones pragmáticas o estéticas: ¿es una *torpeza* mantener la misma etiqueta para esta rama y para la rama progenitora? ¿Sería *engañoso*, por una razón u otra, decir que la rama derecha más que la rama izquierda es la nueva especie[25]?

Nos enfrentamos con el mismo tipo de dificultad cuando tratamos de completar la tarea de codificar con el color rojo la totalidad de la especie aplicando la tinta hacia la base del árbol, para incluir los ancestros de Lulu. No hallaremos ni hiatos ni vinculaciones en este recorrido descendente hasta alcanzar, si persistimos, a los procariotas en la base del árbol. Sin embargo, si al descender también coloreamos hacia los lados, pintamos de rojo primos, tías y tíos de Lulu y sus ancestros, y entonces coloreamos hacia arriba desde estas extensiones laterales, recubriremos eventualmente de rojo la totalidad de la rama en la que reside Lulu, hasta el punto en que al pintar cualquier nódulo inferior, es decir, más temprano en la línea del tiempo evolutivo (por ejemplo, a nivel de A en la figura 4.6), se produce una «filtración» del color rojo en las ramas vecinas que claramente pertenecen a otra especie.

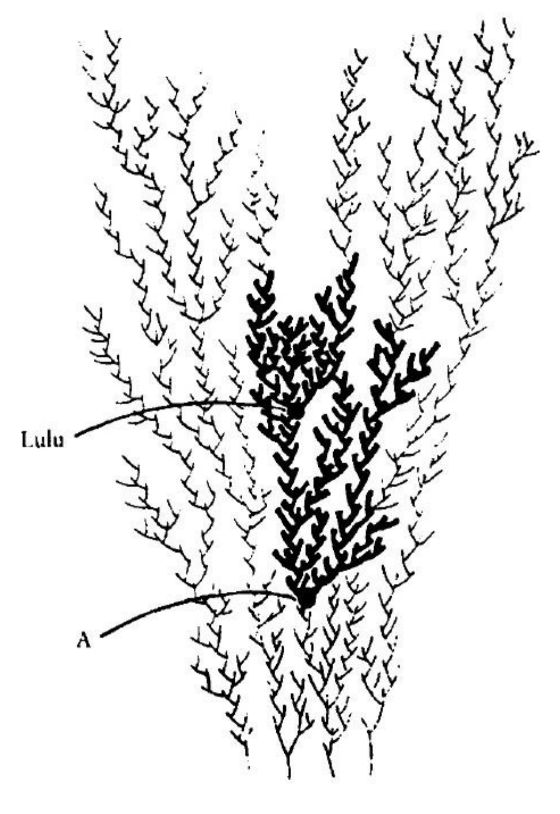


Figura 4.6

Si nos detenemos ahí, podremos estar seguros de que *sólo* se han pintado los miembros de la especie Lulu. Sería discutible que hayamos pasado por alto algunos miembros que merecerían ser coloreados, pero sólo discutible, ya que no existen hechos ocultos ni esencias que puedan decidir la cuestión. Como Darwin señaló, si no fuera por las separaciones que el tiempo y la extinción de los pasos evolutivos intermedios han creado, a pesar de que podemos considerar las formas de vida en una «disposición natural» (de descenso), no podríamos colocarlos en una «clasificación natural»; necesitamos los mayores hiatos posibles entre las formas «supervivientes» para formar los «límites» de tales clases.

El concepto teórico de especie que precede en el tiempo a la teoría de Darwin contiene dos ideas fundamentales: que los miembros de las especies tienen diferentes esencias y que, «por lo tanto», no pueden entrecruzarse. Lo que nosotros hemos deducido posteriormente es que, en principio, podía haber dos subpoblaciones que fueran diferentes *tan sólo* en que sus parejas eran estériles debido a una pequeña incompatibilidad genética. ¿Serían éstas especies diferentes? Podrían parecer idénticas, comer igual, vivir juntas en el mismo nicho ecológico y ser genéticamente, muy similares, pero

reproductivamente aisladas. No serían lo bastante diferentes para ser consideradas como variedades dominantes, pero cumplirían la primera condición para ser dos especies distintas. En efecto, hay casos de «especies hermanas crípticas» que se aproximan a este extremo. Como ya hemos señalado, en el otro extremo tenemos los perros fácilmente distinguibles a simple vista en tipos morfológicos, adaptados a entornos muy diferentes, aunque no aislados reproductivamente. ¿Dónde debemos trazar la línea de separación? Darwin demostró que no necesitamos trazar la línea desde una perspectiva esencialista para seguir avanzando científicamente. Disponemos de la mejor de las razones para darnos cuenta de que estos extremos son improbables: en general, donde existe especiación genética, también existe una diferencia morfológica marcada, o bien una marcada diferencia en la distribución geográfica o (lo que es más probable) ambas cosas. Si esta generalización no fuera verdadera, el concepto de especie no sería importante, pero no necesitamos preguntar exactamente cuánta diferencia (además del aislamiento reproductivo) *es esencial* para que se trate de una diferencia *real* entre especies[26].

Darwin nos demostró que preguntas tales como «¿Cuál es la diferencia entre una variedad y una especie?» son como la pregunta «¿Cuál es la diferencia entre una península y una isla?»[27]. Supongamos que vemos una isla a media milla de la costa con la marea alta. Si pudiéramos caminar hasta allí con marea baja sin mojarnos los pies, ¿sería todavía una isla? Si construimos un puente hasta allí, ¿dejaría de ser una isla? ¿Qué sucedería si construyésemos un sólido paso elevado sobre el agua? Si construimos un canal a través del istmo de una península (como el canal de Cape Cod), ¿la convertimos en una isla? ¿Qué pasaría si un huracán hace el trabajo de excavación en la península? Este tipo de indagación es habitual entre los filósofos. Es la actividad socrática del traficante de definiciones o del cazador de esencias: investigar cuáles son las «condiciones necesarias y suficientes» para ser una X. A veces, casi todo el mundo puede darse cuenta de la falta de sentido de la pregunta, ya que las islas, obviamente, no poseen esencias reales, sino, como mucho, esencias nominales. Pero, en otras ocasiones, puede tratarse de una seria cuestión científica que necesita respuesta.

Transcurrido más de un siglo desde Darwin, persisten todavía serios debates entre los biólogos (más aún entre los filósofos de la biología) sobre cómo definir la *especie*. ¿Deben los científicos definir estos términos? Sí, naturalmente, pero sólo hasta cierto punto. Sucede que hay diferentes conceptos de especie con diferentes usos en biología —el que utilizan los

paleontólogos no es de mucha utilidad para los ecologistas— y no una idea clara acerca de cómo unirlas o clasificarlas por orden de importancia hasta coronar una de estas definiciones (la más importante) como *el* concepto de especie. Por estas razones, me inclino a interpretar los persistentes debates más bien como un vestigio del sentido del orden aristotélico que como una característica cualidad disciplinaria. (Todo esto es controvertido, pero véase Kitcher 1984 y G. C. Williams 1992 para más apoyo y argumentación coincidente, y la reciente antología del tema, Ereshefsky 1992 y Sterelny 1994, un perspicaz ensayo que revisa esta antología).

3. *Coronaciones retrospectivas: la Eva Mitocondrial y la invisibilidad de los inicios*

Cuando tratamos de comprobar si los descendientes de Lulu se separan en más de una especie, debemos mirar para ver si aparece alguna rama gruesa y entonces volver hacia atrás si pensamos que en algún lugar a lo largo de la línea del tiempo evolutivo ha tenido lugar una especiación. Nunca planteamos la sospechosa e importante cuestión de cuándo exactamente se dice que ha ocurrido la formación de la especie. La especiación debe ser considerada como un fenómeno de la naturaleza que posee una curiosa propiedad: ¡no se puede decir qué está ocurriendo durante el tiempo en que ocurre! ¡Sólo mucho más tarde podemos decir qué ha ocurrido, coronando retrospectivamente un acontecimiento cuando se descubre que sus resultados tienen una cierta propiedad. La cuestión no es nuestra limitación epistemológica, como si pudiéramos ser capaces de decir cuándo comienza la separación de una especie tan sólo por disponer de mejores microscopios o si consiguiéramos una máquina del tiempo que nos permitiera volver atrás para observar el momento apropiado. La cuestión es la propiedad objetiva de ser un proceso de especiación. No es una propiedad que un acontecimiento tiene simplemente en virtud de sus propiedades locales espacio-temporales.

Otros conceptos muestran curiosidades similares. En cierta ocasión leía una novela histórica, tan mala que resultaba cómica, en la que un médico francés llega a su casa para cenar una noche del año 1802 y le dice a su mujer: «¿Adivinas lo que he hecho hoy? ¡He asistido al nacimiento de Victor Hugo!». ¿Qué hay de erróneo en esta historia? O bien, consideremos la propiedad de ser una viuda. Una mujer en la ciudad de Nueva York puede adquirir súbitamente esta condición en virtud de los efectos de una bala que

ha alcanzado el cerebro de un hombre en Dodge City, a unas mil millas de distancia. (En los días del salvaje oeste había un revólver denominado el «fabricante de viudas». Si un revólver particular cumple con su calificativo en una ocasión concreta puede tratarse de un hecho que no puede ser establecido por una investigación local espacio-temporal de sus efectos). Este caso plantea la curiosa capacidad para saltar, a través del espacio y del tiempo, desde la convencional naturaleza de la relación del matrimonio, en el que un hecho histórico pasado, una boda, se piensa que ha creado una relación permanente —una relación *formal*— de interés a pesar de los consiguientes extravíos y concretas desgracias (la pérdida accidental de un anillo, la destrucción del certificado de matrimonio, por ejemplo).

La sistematicidad de la reproducción genética no es convencional sino natural, pero esta intensa sistematicidad nos permite pensar *formalmente* acerca de cadenas de causas que se extienden a lo largo de millones de años, cadenas causales que sería, en otras circunstancias, virtualmente imposible de designar, hacer referencia o realizar su seguimiento. Esto nos permite llegar a interesarnos por ellas y razonar rigurosamente, incluso con relaciones más distantes y localmente invisibles que la formal relación de un matrimonio. La especiación es, como el matrimonio, un concepto anclado dentro de un sistema de pensamiento coherente, aunque, a diferencia del matrimonio, no posee características convencionales —bodas, anillos, certificados— por las que pueda ser observado. Este fenómeno de la especiación podemos observarlo con mayor claridad si fijamos nuestra atención en otro ejemplo de coronación retrospectiva, como es la concesión del título de Eva Mitocondrial.

Eva Mitocondrial es la mujer que aparece como el más directo y reciente ancestro, en la línea femenina, de todo ser humano vivo en la actualidad. Cuesta mucho esfuerzo y tiempo imaginarse a esta mujer singular, por lo que conviene revisar en primer lugar el razonamiento que vamos a seguir. Consideremos el grupo A, formado por todos los seres humanos hoy vivos. Cada uno ha nacido de una única madre, así que consideremos el siguiente grupo, el B, constituido por todas las madres de aquellos que viven hoy. El grupo B es necesariamente más reducido que el grupo A, dado que ninguno de los seres humanos vivos hoy tiene más de una madre y algunas madres tienen más de un hijo. Continuemos con el grupo C, el de las madres de todas aquellas madres incluidas en el grupo B. Es todavía más pequeño. Continuemos con el grupo D y el E, y así sucesivamente. Los grupos deben hacerse cada vez más pequeños cuando retrocedemos en cada generación.

Obsérvese que cuando estamos retrocediendo a través de los años, quedan excluidas muchas mujeres que fueron contemporáneas de las ubicadas en nuestro grupo. Entre las mujeres excluidas se encuentran aquellas que habían vivido y fallecido sin tener hijos. Eventualmente, este grupo debe estrecharse hasta una mujer: la mujer que está más cerca de ser el directo y más cercano antepasado femenino de todos los que hoy viven en la Tierra. ¡Ella es la Eva Mitocondrial!, denominada así (Cann y otros 1987) porque como las mitocondrias de nuestras células pasan solamente a través de la línea materna, todas las mitocondrias de todas las células de toda la gente hoy viva son descendientes directas de la mitocondria de las células de esta Eva.

El mismo argumento lógico establece que hay también —debe haber— un Adán: el que está más cerca de ser el directo ancestro de todos los que viven hoy. Podemos llamarlo Adán Cromosoma-Y, ya que todos nuestros cromosomas Y pasan a través de la línea paterna, del mismo modo que nuestras mitocondrias pasan a través de la línea —materna[28]. ¿Era el Adán Cromosoma-Y el esposo o amante de la Eva Mitocondrial? La respuesta es no, casi con toda certeza. Existe tan sólo una pequeñísima probabilidad de que estos dos individuos vivieran al mismo tiempo. (Siendo la paternidad un negocio que consume mucho menos tiempo y energía que la maternidad, es *lógicamente* posible que Adán Cromosoma-Y viviera muy recientemente y estuviera muy, pero que muy ocupado, en el dormitorio, haciéndole morder el polvo a Errol Flynn. Este Adán pudo, en principio, ser el tata tatarabuelo de todos nosotros. Esta versión es tan poco probable como que el Adán Cromosoma-Y y la Eva Mitocondrial formaran una pareja).

La Eva Mitocondrial ha conseguido recientemente titulares en las noticias debido a que los científicos que la han bautizado piensan que pueden analizar los patrones del ADN de las mitocondrias de las diferentes personas vivas hoy y deducir, con este método, cuándo vivió la Eva Mitocondrial, si ha sido hace poco e incluso dónde vivió. De acuerdo con sus cálculos originales, la Eva Mitocondrial vivió en África hace relativamente poco, menos de trescientos mil años e incluso puede que menos de la mitad de este tiempo. Sin embargo, estos métodos de análisis son controvertidos y la hipótesis de la Eva Mitocondrial puede haber quedado afectada a la fuerza. Deducir *dónde* y *cuándo* es una tarea mucho más dificultosa que deducir *que* existió una Eva Mitocondrial, algo que nadie niega. Veamos algunas de las pocas cosas que ya conocemos acerca de la Eva Mitocondrial, dejando a un lado las controversias. Sabemos que tuvo al menos dos hijas de las que habría sobrevivido descendencia. (Si hubiese tenido sólo, una hija, ésta habría

heredado la corona de Eva Mitocondrial). La llamaremos Amy, para distinguir su título de su nombre propio. Amy utiliza el título de Eva Mitocondrial, por haber sido justamente la madre fundadora de la línea actual de los seres humanos[29]. Es importante que recordemos que *desde todos los puntos de vista*, no hay probablemente nada llamativo o especial acerca de la Eva Mitocondrial; ella no fue con seguridad la primera mujer o la fundadora de la especie del *Homo sapiens*. Hubo antes muchas otras mujeres de nuestra especie, pero sucedió que no tuvieron descendientes en línea directa que condujeran a los seres humanos vivos hoy. También es verdad que la Eva Mitocondrial no sería con toda probabilidad la más fuerte, la más rápida, la más bella y la más fecunda entre las otras mujeres de su tiempo.

Para resaltar hasta qué punto no había nada especial en la Eva Mitocondrial —es decir, Amy— supongamos que mañana, miles de generaciones después, un nuevo virus muy violento se extiende por toda la Tierra azotando al 99 por ciento de la raza humana en unos pocos años. Los supervivientes, los afortunados que posean alguna resistencia innata al virus, estarán probablemente muy relacionados por parentesco. *Su* ancestro femenino más cercano, común y directo —lo llamaremos Betty— sería alguna mujer que vivió cientos o miles de generaciones más tarde que Amy, y la corona de Eva Mitocondrial pasaría a ella, retroactivamente. Betty pudo haber sido la fuente de la mutación que, siglos más tarde, se convirtió en salvadora de su especie, aunque esta mutación no le produjo a *ella* ningún bien, dado que el virus contra el que había triunfado no existía por aquel entonces. La cuestión es que la Eva Mitocondrial sólo puede ser coronada *retrospectivamente*. Este histórico papel de pivote es determinado no tan sólo por los accidentes acaecidos en el tiempo de Amy, sino también por los accidentes ocurridos en tiempos posteriores. ¡Hablemos de contingencia masiva! Si el tío de Amy no la hubiera salvado de morir ahogada cuando tenía tres años, ninguno de *nosotros* (con nuestro particular ADN en las mitocondrias, gracias precisamente a Amy) hubiera existido. Si las nietas de Amy hubieran muerto de hambre en la infancia —como muchos niños murieron en aquellos días—, tampoco hubiéramos existido.

La curiosa invisibilidad de la corona de la Eva Mitocondrial durante su vida es más fácil de comprender y aceptar que la casi invisibilidad del comienzo de todas las especies. Si las especies no son eternas, entonces todo el tiempo puede ser dividido, de algún modo, en el tiempo antes de la existencia de la especie *x* y en todos los tiempos subsiguientes. Pero ¿qué es lo que ha sucedido en la interfaz? Puede servirnos de ayuda pensar en un

rompecabezas similar que ha desconcertado a mucha gente. ¿Se ha preguntado alguna vez cuando ha escuchado un nuevo chiste, de dónde procede? Si es como la mayoría de la gente que yo he conocido o de la que tengo noticia, nunca hace chistes; sino que transmite, acaso con «mejoras», algo que ha oído de alguien, que a su vez se lo escuchó de alguien, que… Ahora, sabemos que el proceso no puede seguir así siempre. Un chiste sobre el presidente Clinton, por ejemplo, no puede tener más de un año, poco más o menos. Pero ¿quién hace los chistes? Los autores de los chistes (como contraste con los proveedores de chistes) son invisibles[30]. Nadie parece haber asistido al acto de su autoría. Incluso existe la creencia —una «leyenda urbana»— de que todos los chistes son inventados en las prisiones por reclusos, esa gente peligrosa y antinatural, al contrario que el resto de nosotros, que no tienen nada mejor que hacer con su tiempo que inventar los chistes de moda en sus celdas secretas y *underground*. Pura estupidez. Es difícil de creer —pero debe ser verdad— que los chistes que oímos y contamos han evolucionado a partir de historias previas, sometidos a adiciones y revisiones, cuando pasan de unos a otros. Lo característico de los chistes es que no tienen autor; su autoría se distribuye entre docenas, cientos o miles de contadores de chistes, consolidándose durante algún tiempo en alguna versión divertida, especialmente tópica y actual, antes de convertirse en chiste durmiente, como los ancestros a partir de los cuales creció. Por la misma razón, es difícil ser testigo de la especiación.

¿Cuándo se origina una especie? En muchos casos (quizá la mayoría, quizá casi todos; los biólogos no se ponen de acuerdo acerca de lo importante que son las excepciones) la especiación depende de una separación geográfica en la que un pequeño grupo —puede ser una simple pareja, macho y hembra — se aparta y comienza una línea que queda aislada desde el punto de vista reproductivo. Esta es una especiación *alopátrica*, en contraste con la formación *simpátrica*, que no implica la existencia de barreras geográficas. Supongamos que podemos observar la partida y el reasentamiento del grupo fundador. El tiempo pasa y varias generaciones se suceden. ¿Ha tenido lugar la formación de una nueva especie? Todavía no, ciertamente. No sabremos hasta pasadas muchas generaciones si estos individuos deben ser coronados o no como iniciadores de una especie.

No es y *no puede ser* algo interno o intrínseco de los individuos —e incluso de los individuos en cuanto que se adaptan lo más posible a su ambiente— de lo que se deduzca que ellos fueron —como ellos más tarde llegaron a ser— los fundadores de una nueva especie. Podemos imaginarnos,

si lo deseamos, un caso extremo (e improbable), en el que una simple mutación pueda garantizar el aislamiento reproductivo en una sola generación, aunque naturalmente, si el individuo que ha sufrido la mutación cuenta o no como un fundador de especies o simplemente como un fracaso de la naturaleza, no depende de su estructura como individuo o biografía, sino de lo que suceda en las subsiguientes generaciones —si hay alguna— de descendientes.

Darwin no fue capaz de presentar en *El origen de las especies*, un solo ejemplo de especiación mediante la selección natural. Su estrategia en ese libro fue desarrollar con detalles la evidencia de que la selección artificial realizada por los criadores de perros y de palomas podía llegar a generar grandes diferencias a través de una secuencia de cambios graduales. Darwin puntualizó que la elección *deliberada* por los criadores no era esencial; los ejemplares enanos de la camada tendían a no ser valorados y, por lo tanto, a no reproducirse tanto como sus hermanos más valorados; de este modo, los criadores de animales, sin ninguna política consciente de cruzamientos, presidían inconscientemente un proceso seguro de revisión de diseño. Darwin ofreció el bello ejemplo del perro *King Charles spaniel*, «que fue inconscientemente modificado en gran parte desde el tiempo del monarca» (*El origen de las especies*), como puede confirmarse si se observa con detenimiento a los perros en varios retratos del rey Carlos. Darwin consideró estos casos como las consecuencias de una «selección inconsciente» realizada por los criadores, y la utilizó como un persuasivo puente que sus lectores atravesarían para alcanzar la hipótesis de una selección aún más inconsciente, relacionada con el impersonal medio ambiente. Pero Darwin tuvo que admitir, cuando se cuestionaba su tesis, que no podía presentar casos de criadores de animales que hubiesen producido una nueva especie. Tales cruzamientos habían producido definitivamente diferentes *variedades* pero ninguna nueva especie. Los perros Dachshund y San Bernardo no eran diferentes especies aunque sí lo fueran en apariencia. Darwin cedió demasiado en este punto, ya que pudo argumentar, de forma totalmente correcta, que era demasiado pronto para decir si había o no algunos ejemplos de nuevas especies producidas por selección artificial. En el futuro, se podría haber descubierto cualquier perrito faldero como *habiendo sido* el miembro fundador de una especie que se separó del *Canis familiaris*.

La misma moraleja se aplica, naturalmente, a la creación de nuevos géneros, familias e incluso reinos. La mayor ramificación que podemos coronar retrospectivamente, como es la separación entre las plantas y los

animales, comienza por una segregación de dos *depósitos* de genes, cada parte tan inescrutable y poco llamativa en su tiempo como las separaciones transitorias de miembros de una misma población.

4. *Modelos, hipersimplificación y explicación*

Mucho más interesante que la cuestión de cómo marcar los límites de las especies son todas aquellas otras que se ocupan de las formas de las ramas; y aún todavía más interesante, de las formas de los espacios vacíos entre las ramas. ¿Qué tendencias, fuerzas y principios —o acontecimientos históricos— han influido en estas formas o las han hecho posible? Los ojos han evolucionado de modo independiente en docenas de líneas evolutivas, pero las plumas probablemente sólo en una. Como ha subrayado John Maynard Smith, los mamíferos consiguieron cuernos, pero no los pájaros. «¿Por qué el modelo de variación debe ser limitado a una vía? La respuesta es sencilla: no lo sabemos» (Maynard Smith 1986:41).

No *podemos* rebobinar la cinta de la vida y proyectarla de nuevo para observar lo que sucedió, de modo que el único camino para contestar a las preguntas acerca de tan inmensos, y experimentalmente inaccesibles, modelos es lanzarse de forma temeraria al vacío con la arriesgada táctica de la deliberada hipersimplificación. Esta táctica tiene una larga y distinguida historia en la ciencia, aunque tiende a provocar controversias, dado que los científicos poseen diferentes umbrales, traspasados los cuales se ponen nerviosos si los detalles recalcitrantes son tratados con rapidez y no excesivo rigor. La física de Newton fue sobrepasada por Einstein pero todavía es una buena aproximación para casi todos los propósitos. Ningún físico puso objeciones cuando la NASA utilizó la física de Newton para el cálculo de fuerzas en el lanzamiento y en la trayectoria de la órbita de la lanzadera espacial, aunque, hablando en serio, este es un caso de uso deliberado de una falsa teoría para hacer posible los cálculos. Con el mismo espíritu —se dice—, los fisiólogos que estudian los mecanismos del cambio del ritmo metabólico tratan, por lo general, de evitar las extrañas complejidades de la física cuántica subatómica, dando por supuesto que todos los efectos cuánticos serán anulados o, por el contrario, estarán por debajo del umbral de sus modelos. Es habitual que esta táctica proporcione buenos resultados, aunque nunca se puede estar seguro, ya que lo que es una complicación molesta en un científico, en otro se convierte en la clave del misterio. Y esto puede funcionar también en sentido contrario: la clave se suele descubrir

saliendo de las trincheras y optando por una visión panorámica del campo de batalla.

En cierta ocasión mantuve un debate con Francis Crick acerca de las virtudes y los vicios del conexionismo, un movimiento dentro de la ciencia cognitiva que diseña fenómenos psicológicos mediante la construcción de modelos de las fuerzas de conexión entre los nodos, como «redes neurales» simuladas en los computadores, de una manera nada realista y desde luego extraordinariamente simplificada. «¡Estas gentes pueden ser buenos ingenieros —afirmó Crick (al menos así lo recuerdo)—, pero lo que están haciendo es una ciencia terrible! Esa gente vuelve voluntariamente la espalda a lo que ya conocemos sobre las interacciones de las neuronas, de modo que sus modelos son del todo inútiles como modelos de la función cerebral». Esta crítica me sorprendió, ya que Crick es famoso por su brillante oportunismo al descubrir la estructura del ADN; mientras que otros luchaban a través del recto y estrecho camino de la construcción estricta desde la evidencia, él y Watson se decidieron, con audacia y optimismo, por seguir un camino lateral a través del cual alcanzaron resultados gratificantes. En cualquier caso, sentí la curiosidad de conocer la amplitud de la denuncia que había lanzado Crick. ¿Habría dicho lo mismo a los que se ocupan de la genética de las poblaciones? El término despreciativo que se da a algunos de los modelos utilizados es «genética de saquito de alubias», ya que pretenden que los genes de uno y de otro son algo así como numerosos abalorios codificados mediante colores, ensartados en un hilo. Lo que ellos llaman gen (o un *alelo* en un *locus*) tiene sólo una semejanza pasajera con la intrincada maquinaria de las secuencias de un codon en las moléculas de ADN. Pero gracias a esta deliberada simplificación, sus modelos pueden tratarse mediante cálculos con la ayuda del ordenador, haciendo posible descubrir y confirmar muchos modelos a gran escala en el fluir de los genes, que de otra manera serían del todo invisibles. Si se añaden complicaciones al modelo su investigación tendería a bloquearse. Pero ¿su investigación es una buena ciencia? Crick replicó que él mismo había reflexionado sobre la comparación y que tenía que decir a los aplicados a la genética de las poblaciones que ¡eso no era ciencia!

Mis gustos en ciencia son más indulgentes y quizá debería esperarse esta postura de un filósofo, aunque tengo mis razones: creo que el argumento es sólido y que no sólo los modelos «hiper» simplificados *explican* de verdad lo que necesitamos explicar, sino que modelos más complicados no pueden hacer el trabajo. Cuando lo que provoca nuestra curiosidad son los *grandes modelos* de los fenómenos, necesitamos una explicación a escala adecuada.

En muchos casos esto es obvio. Si se desea conocer por qué el atasco del tráfico en una ciudad tiende a producirse a cierta hora todos los días, quedaremos desconcertados después de haber reconstruido cuidadosamente el proceso de conducción, frenado y aceleración de los miles de conductores cuyas diferentes trayectorias se han sumado para provocar aquellos atascos.

Imaginemos los trazados que siguen todos los electrones en una calculadora manual cuando multiplica dos números y produce la respuesta correcta. Se puede estar seguro al 100 por ciento de que en el proceso se comprenden cada uno de los millones de micropasos causales y todavía quedaremos absolutamente desconcertados acerca del *por qué* e incluso el *cómo* siempre se consigue la respuesta *correcta* a las cuestiones planteadas. Si esto no es obvio, imaginemos que alguien fabrica una calculadora —desde luego una costosa travesura— que suele producir respuestas erróneas. Esta última obedecería a las mismas leyes físicas que la buena calculadora y funcionaría a través de los mismos microprocesos. Podríamos conseguir explicaciones *perfectas* sobre cómo ambas calculadoras trabajan a nivel electrónico y seguir siendo absolutamente incapaces de explicar el muy interesante hecho de que una calculadora da respuestas correctas y la otra erróneas. Este es el tipo de casos que muestra lo que hay de ridículo en las formas descabelladas del reduccionismo; naturalmente, no se pueden explicar todos los modelos que interesan a nivel de la física (o de la química o cualquier otro nivel más bajo). Es verdaderamente innegable para todos aquellos fenómenos cotidianos, como el atasco del tráfico y la calculadora de bolsillo; cabe esperar que también sea verdad para los fenómenos biológicos. (Para más detalles, véase Dennett 1991b).

Consideremos ahora una cuestión paralela en biología, extraída de un libro de texto estándar: ¿por qué las jirafas tienen el cuello largo? Hay una respuesta que puede, en principio, ser «leída» en el propio árbol de la vida si le echamos una mirada: cada jirafa tiene un cuello de esa longitud porque sus padres tienen cuellos de esa misma longitud, y así podemos seguir retrocediendo a través de generaciones. Si comprobamos, uno por uno, los cuellos de las jirafas, veremos que el cuello de cada jirafa viva ha sido seguido a través de ancestros de cuello largo todo el recorrido hacia atrás… hasta ancestros que ya no tenían cuello. Así es como las jirafas han llegado a tener cuellos largos. Fin de la explicación (y si esta explicación no le satisface, aún quedará menos satisfecho si la respuesta entra en detalles acerca del desarrollo individual y de la historia nutricional de cada jirafa en su propio linaje).

Cualquier explicación aceptable de los modelos que observamos en el árbol de la vida debe mostrar las diferencias: ¿por qué nos fijamos en este modelo real más que en otro modelo, o en ninguno? ¿Cuáles son las alternativas que no se han hecho reales y que necesitan ser consideradas, y cómo están organizadas? Para responder a estas preguntas necesitamos ser capaces de discutir acerca de lo que es posible además de lo que es real.

CAPÍTULO 5
Lo posible y lo real

I. *¿Existen grados de posibilidad?*

> Aunque existen muchas vías por las que llegar a ser un ser vivo, es cierto que hay muchas más vías para estar muerto o, mejor dicho, no vivo.
>
> Richard Dawkins, *El relojero ciego*

> Cualquier forma particular de vida no existente puede deber su ausencia a una de estas dos razones. Una es la selección negativa, la otra es que nunca han aparecido las mutaciones necesarias.
>
> Mark Ridley, *The Problems of Evolution*

> Imaginemos, por ejemplo, un posible hombre obeso en un portal; y de nuevo un posible hombre calvo en ese portal. ¿Son ellos el mismo hombre posible o se trata de dos hombres posibles? ¿Cómo lo decidimos? ¿Cuántos hombres posibles hay en ese portal? ¿Hay más hombres posibles delgados que obesos? ¿Cuántos de ellos son similares? ¿El parecido entre ellos les haría ser uno solo? ¿No son similares dos cosas posibles? ¿Es esto lo mismo que decir que es imposible para dos cosas ser similares? Y, finalmente, ¿es el concepto de identidad simplemente inaplicable a posibles no reales?
>
> Willard Van Orman Quine, «*Ow What There Is*»

Al parecer, existen al menos cuatro clases o grados diferentes de posibilidad: lógica, física, biológica e histórica, por este orden. La más indulgente es la mera posibilidad lógica, que, de acuerdo con la tradición filosófica, sólo exige ser descriptible sin contradicción. Superman, que vuela más rápido que la velocidad de la luz, es *lógicamente* posible, pero Superman, que vuela más rápido que la velocidad de la luz aunque *sin moverse hacia ninguna parte*, no es lógicamente posible. Sin embargo. Superman no es *físicamente* posible, dado que una ley física establece que nada puede moverse más rápido que la velocidad de la luz. No son pocas las dificultades que plantea esta superficial distinción. ¿Cómo podemos distinguir las leyes de la física fundamentales de

las leyes de la lógica? ¿Es física o lógicamente imposible, por ejemplo, viajar hacia atrás en el tiempo? ¿Cómo podemos decir con seguridad si una descripción que es *aparentemente* coherente —como la historia de la película *Regreso al futuro*— es contradictoria en sí misma o sencillamente niega una fundamental (pero no necesariamente lógica) tesis de la física? No son pocas, tampoco, las aportaciones de la filosofía al tratamiento de estas dificultades, así es que deberemos conocerlas y pasar al siguiente grado.

Superman vuela sencillamente lanzándose al espacio y adoptando en el aire una garbosa postura, un talento que, en realidad, es físicamente imposible. ¿Es posible, de acuerdo con las leyes de la física, un caballo volador? El modelo estándar de la mitología nunca hubiera despegado del suelo —un hecho en el ámbito de la física (aerodinámica), no de la biología— pero un caballo con alas adecuadas podría presumiblemente permanecer en vuelo. Tendría que ser un caballo pequeño, algo que los ingenieros aeronáuticos podrían calcular teniendo en cuenta la *ratio* peso-solidez, la densidad del aire y otros factores. Pero cuando descendemos al tercer grado de la posibilidad, la *posibilidad biológica*, para lo que comenzamos analizando la resistencia de los huesos y los requerimientos de carga útil para mantener en funcionamiento la maquinaria voladora, nos preocupamos del desarrollo y el crecimiento, de su metabolismo y de otros fenómenos claramente biológicos. Incluso podemos deducir que los caballos voladores son posibles dado que los murciélagos son reales. Podría ser que incluso caballos voladores de tamaño natural fueran biológicamente posibles si se tiene en cuenta que, en algún tiempo, los reptiles voladores («Pteranodontos»), y otras criaturas voladoras, se aproximaban a este tamaño. No hay nada que supere la realidad, presente o pasada, desde la posibilidad confirmada. Cualquier cosa que es o ha sido real es obviamente posible. ¿No es así?

Las lecciones sobre realidad son difíciles de interpretar. ¿Podrían ser tales caballos voladores realmente viables? ¿Necesitarían, acaso, ser carnívoros para almacenar la suficiente energía que les permitiera elevarse y volar? Quizás —a pesar de que los murciélagos se alimentan de frutas—, sólo un caballo carnívoro podría despegar de la tierra. ¿Sería posible un caballo carnívoro? Quizás un caballo carnívoro sería biológicamente posible *si pudiera evolucionar*, pero ¿tal cambio de dieta sería accesible allí donde los caballos tendrían que comenzar la evolución? Y, dejando a un lado la posibilidad de una radical cirugía constructiva, ¿un descendiente de un caballo podría tener, al mismo tiempo, extremidades y alas? Los murciélagos,

después de todo, han convertido sus extremidades delanteras (brazos) en alas. ¿Existe alguna historia evolutiva de revisión esquelética que haya generado un mamífero de seis extremidades?

Esto nos lleva a nuestro cuarto grado de posibilidad, la *posibilidad histórica*. Pudo haber existido un tiempo, en un pasado muy remoto, en que la posibilidad de un mamífero con seis extremidades sobre la tierra no hubiera sido todavía excluida por adelantado, pero también puede ser verdad que una vez que nuestros ancestros con cuatro extremidades fueron seleccionados para moverse por la tierra, la arquitectura básica de cuatro extremidades estaba tan profundamente anclada en nuestros esquemas de desarrollo que, en ese tiempo, una modificación *ya no era posible*. Pero incluso esta distinción podría no ser tajante. ¿Esta modificación del plan fundamental de construcción del todo imposible o muy improbable es tan resistente al cambio que sólo una improbable secuencia de golpes selectivos podría dar lugar a su existencia? Parece que puede haber dos clases o grados de imposibilidad biológica: la violación de una *ley biológica de la naturaleza* (si existe alguna) y la «mera» consignación biohistórica del olvido.

La imposibilidad histórica es simplemente un asunto de oportunidades pasadas por alto. Hubo un tiempo en que a muchos de nosotros nos preocupaba la posibilidad de que Barry Goldwater llegara a ser elegido presidente de Estados Unidos, pero esto no sucedió, y después de 1964, las probabilidades contrarias a tal eventualidad parecen haberse incrementado hasta niveles tranquilizadores. Cuando los billetes de lotería se ponen a la venta, crean una oportunidad en nosotros: podemos elegir comprar un billete, si esta adquisición se hace para una cierta fecha. Si compramos un billete se crea una oportunidad nueva —la oportunidad de ganar—, pero pronto esta oportunidad se desliza hacia el pasado, por lo que ya no es posible ganar *aquellos* millones de dólares. ¿Es una ilusión cotidiana creer que tenemos oportunidades, oportunidades *reales*? ¿En qué sentido *podíamos* haber ganado? ¿Qué diferencia existe entre que el número del billete ganador de la lotería sea extraído *después* de que hayamos comprado nuestro billete, o tenemos aún una oportunidad de ganar, una oportunidad real, si el número ganador es guardado en un cofre sellado antes de que los billetes sean puestos a la venta (Dennett 1984)? ¿Existe *todavía* una oportunidad real? ¿Puede suceder alguna otra cosa que lo que realmente sucede? Esta terrible hipótesis, la idea de que sólo lo real es posible, ha sido llamada *actualismo* (A. J. Ayer 1968). Es generalmente desdeñada, por buenas razones, pero estas razones

son raramente discutidas. (Dennett 1984 y Lewis 1986:36-38, ofrecen buenas razones para abandonar el actualismo).

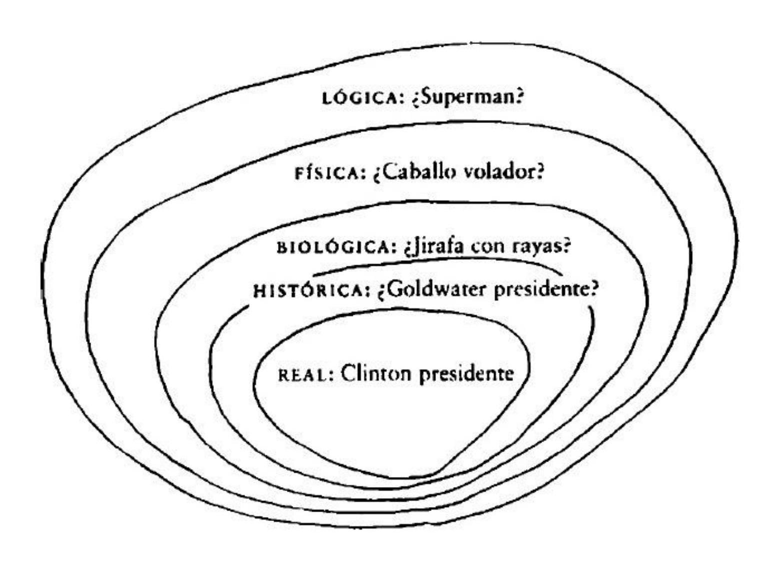


Figura 5-1

Estas ideas familiares y *prima facie* fiables acerca de la posibilidad pueden representarse en un diagrama, aunque cada frontera en esta figura es discutible. Como sugieren las preguntas de Quine, reproducidas en la cabecera de este capítulo, hay algo inverosímil en las categorías causales de los objetos meramente posibles, pero dado que la ciencia no puede aún expresar —sólo confirmar— los tipos de explicaciones que deseamos sin dibujar tal distinción, es poco probable que podamos renunciar simplemente a toda esta discusión. Cuando los biólogos se preguntan si es posible un pájaro con cuernos —o incluso una jirafa con rayas en lugar de manchas—, las preguntas que están planteando resumen lo que deseamos que la biología nos descubra. Alertados por Quine, podemos quedar atrapados en las dudosas implicaciones metafísicas de la afirmación de Richard Dawkins respecto a que hay muchas más maneras de morir que de estar vivo, aunque él se está aproximando a algo importante. Trataremos de encontrar una vía que nos permita refundir tales afirmaciones en una estructura metafísica más modesta y menos contenciosa; Darwin, al empezar en la zona media del proceso evolutivo, nos dio precisamente el punto de apoyo que necesitamos. *En primer lugar*, podemos ocuparnos de la relación entre las posibilidades históricas y las biológicas y entonces, quizás, estas reflexiones nos sugerirán algunas ideas sobre cómo dar sentido a las variedades mayores (la lógica y la física)[31].

2. *La Biblioteca de Mendel*

Al poeta argentino Jorge Luis Borges no se le considera un filósofo, aunque en sus relatos ha proporcionado a la filosofía algunos de los más valiosos experimentos mentales, la mayor parte de ellos reunidos en la asombrosa colección titulada *Ficciones*. Entre lo mejor de estas historias se halla la fantasía —en realidad es más una reflexión filosófica que una narración— que describe la Biblioteca de Babel. Para nosotros, la Biblioteca de Babel será una visión que nos ayudará a responder a preguntas muy difíciles en el ámbito de las posibilidades biológicas, así que haremos una pausa en nuestro discurso para explorarla con alguna extensión. Borges nos relata las desoladoras exploraciones y especulaciones de unas personas que se encuentran viviendo en un enorme almacén de libros, estructurado como un panal de abejas, compuesto de miles (o millones o miles de millones) de huecos hexagonales rodeados por galerías recubiertas con estantes. Si nos apoyamos en una barandilla y miramos hacia arriba o hacia abajo no veremos ni el extremo superior ni el inferior de estos espacios vacíos. Nadie ha encontrado todavía un hueco hexagonal que no esté rodeado por seis huecos vecinos. Ellos se preguntan: ¿es infinita la biblioteca? Eventualmente deciden que no lo es, pero también puede serlo porque parece que en sus estantes —sin ningún orden, ¡qué lástima!— *se encuentran todos los libros posibles*.

Supongamos que cada libro tiene 500 páginas y cada página tiene 40 líneas de 50 espacios, de modo que hay dos mil caracteres por página. Cada espacio o está vacío o tiene un carácter impreso en él, escogido entre un conjunto de 100 (las letras mayúsculas y las letras minúsculas del inglés y de otras lenguas europeas, más los espacios vacíos y las marcas de las puntuaciones)[32]. En algún lugar de la Biblioteca de Babel hay un volumen constituido en su totalidad por páginas en blanco y otro volumen lleno de signos de interrogación, pero la inmensa mayoría consiste en un galimatías tipográfico: ninguna regla ortográfica ni gramatical, y, por descontado, ninguna regla de sentido, prohíbe la inclusión de un volumen. Dos mil caracteres por página, a 500 páginas por libro, suman 1.000.000 de caracteres por libro, así que hay $100^{1.000.000}$ libros en la biblioteca de Babel. Dado que se estima[33] que en la región del universo hay solamente 100^{40} (más o menos) *partículas* (protones, neutrones y electrones) que podemos observar, la biblioteca de Babel no es ni de lejos un objeto físicamente posible, pero gracias a las estrictas reglas con las que Borges la construyó en su imaginación, podemos opinar sobre ella con claridad.

¿Es este verdaderamente el conjunto de *todos* los libros posibles? Obviamente no, ya que han sido restringidos a los impresos «solamente» con 100 caracteres diferentes, excluyendo, suponemos, los caracteres del griego, ruso, chino, japonés y árabe, y por lo tanto, pasando por alto muchos de los más importantes libros *reales*. Naturalmente, la Biblioteca de Babel contiene soberbias traducciones de todos estos libros reales al inglés, francés, alemán, italiano…, así como incontables trillones de malas traducciones de cada libro. Los libros de más de 500 páginas se encuentran allí, empezando con un volumen y siguiendo, sin interrupción, por algún otro volumen o volúmenes.

Es divertido pensar en alguno de los volúmenes que puedan encontrarse en algún lugar de la Biblioteca de Babel, Uno de ellos es la mejor y más precisa biografía que se ha escrito sobre cada uno de nosotros en un volumen de 500 páginas, desde el momento de nuestro nacimiento hasta nuestra muerte. Localizarla, sin embargo, será del todo imposible (¡esta escurridiza palabra!), dado que la Biblioteca también contiene billones y billones de volúmenes que son biografías extraordinariamente detalladas de cuando cumplimos diez, veinte, treinta, cuarenta años de nuestro nacimiento, y completamente falsas en lo que se refiere a los subsiguientes acontecimientos de nuestra vida, según billones de relatos diferentes y divertidos. Pero incluso encontrar un volumen legible en este enorme almacén de libros es extremadamente improbable.

Necesitamos hacer algunas precisiones sobre las cantidades implicadas. La Biblioteca de Babel no es infinita, así que la probabilidad de encontrar en ella algo interesante no es literalmente infinitesimal[34]. Estas palabras, dichas en tono familiar, son exageradas —sorprendimos en Darwin algo semejante en su resumen, donde utilizó un ilícito «infinitamente»— por lo que debemos evitarlas. Por desgracia, todas las metáforas habituales —«astronómicamente grande», «una aguja en un pajar», «una gota en el océano»— se quedan cortas. Ninguna cantidad astronómica *real* (como el número de las partículas elementales del universo o el tiempo transcurrido desde el Big Bang medido en nanosegundos) es siquiera apreciable contra el telón de fondo de estos enormes pero finitos números. Si un volumen legible de la Biblioteca fuera tan fácil de encontrar como una gota en el océano, ¡no habría problema! Si fuéramos arrojados aleatoriamente dentro de la Biblioteca, nuestra probabilidad de encontrar un volumen con algo así como una sentencia gramatical sería tan extraordinariamente evanescente que podríamos escribir el término en mayúsculas —«Evanescentemente pequeña»— y emparejar esta definición con la idea de una probabilidad «muchísimo más que

astronómica», cuyo acrónimo en inglés podría ser VASTLY («Very-much-more-than-AST ronomicaLY»)[35].

Moby Dick está naturalmente en la Biblioteca de Babel, aunque también hay unos 100.000.000 de mutantes impostores que difieren del *Moby Dick* canónico en un solo error tipográfico. Este no es todavía un número «muchísimo más que astronómico», pero el total se incrementa rápidamente cuando añadimos las variantes que difieren por 2 0 10 0 1000 errores tipográficos. Aun en un volumen con 1.000 errores tipográficos —2 páginas de media— sería reconocible sin equivocación como *Moby Dick*, y hay una cifra «muchísimo más que astronómica» de aquellos volúmenes. No habría problemas con estos volúmenes si pudiéramos hallar solamente uno de ellos. Prácticamente todos seguirían siendo una agradable lectura y en todos ellos aparecería la misma historia, con la salvedad de diferencias mínimas, casi indistinguibles. Pero, sin embargo, no ocurriría en todos los casos. A veces, un simple error tipográfico puede ser fatal si se encuentra en una posición crucial. Peter De Vries, otro escritor de libros de ficción con un fantástico enfoque filosófico, publicó una vez una novela que comenzaba[36]: «Call me, Ishmael», 'Llámame, Ismael', en lugar del original «Call me Ishmael», 'Llámame Ismael'. ¡Véase lo que puede hacer una simple coma! O bien, consideremos las diferentes variaciones de la frase inicial de *Moby Dick* que empiezan con: «Ball me Ishmael», 'Báilame Ismael' en lugar de «Call me Ishmael», 'Llámame Ismael'.

En el relato de Borges, los libros no están colocados en los estantes en ninguna clase de orden, pero incluso si los encontrásemos escrupulosamente alfabetizados, tendríamos insolubles problemas para hallar *el* libro que estamos buscando (por ejemplo, la versión «esencial» de *Moby Dick*). Podemos imaginarnos ahora que estamos viajando en una nave espacial a través de la galaxia *Moby Dick* de la Biblioteca de Babel. Esta galaxia es, en sí misma, «muchísimo más que astronómicamente» mayor que la totalidad del universo físico, así que, no importa la dirección que sigamos durante siglos hacia el fin, incluso si pudiéramos viajar a la velocidad de la luz, todo lo que veríamos serían copias virtualmente indistinguibles de *Moby Dick*; nunca encontraríamos nada parecido a otra cosa. *David Copperfield* se encuentra en el espacio a una distancia inimaginable, aun cuando sabemos que hay un camino —un corto camino que ignora billones de billones de otros caminos— que conduce de un gran libro a otro por cambios tipográficos singulares. (Si nos encontramos en este camino, comprobaremos que es casi imposible saber, por inspección local, qué dirección debemos tomar para dirigirnos hacia

David Copperfield, incluso si tenemos a mano ejemplares de los dos libros que son nuestro objetivo).

En otras palabras, el espacio *lógico* es tan vasto astronómicamente que muchas de nuestras habituales ideas acerca de la localización, la búsqueda y el hallazgo de otras actividades mundanas y prácticas, no tienen aplicación directa. Borges colocó los libros en las estanterías en orden aleatorio, un bonito detalle del cual extrajo varias reflexiones sugerentes, pero prestemos atención a los problemas que se habría creado a sí mismo si hubiese intentado colocarlos en orden alfabético en su colmena. Dado que solamente existen unos cien caracteres alfabéticos diferentes (en nuestra versión), podemos organizar algunas secuencias específicas en orden alfabético; por ejemplo, a, A, b, B, c, C… z, Z, ?, ;…,,,.,!,), (, %, …à, â, è, ê, é… De este modo podemos colocar todos los libros que comiencen con la misma letra en el mismo *piso*. Ahora nuestra biblioteca tiene sólo 100 pisos de altura, es decir, es más baja que el World Trade Center. Se puede dividir cada piso en 100 *pasillos*, cada uno de ellos recubierto por filas de libros cuyo segundo carácter tipográfico es el mismo, un pasillo para cada carácter, en orden alfabético. En cada pasillo podemos colocar 100 estantes, uno por cada tres aperturas. Todos los libros que empiezan por «aarvarks love Mozart» («el cerdo hormiguero ama a Mozart») —¡son muchísimos!— están colocados en el mismo estante (el estante «*r*») en el primer pasillo del primer piso. Pero este es un estante extraordinariamente largo, así que quizá sería mejor que colocáramos los libros en cajones archivadores, en ángulo recto con el estante, un cajón para cada cuarta letra. De este modo, cada estante puede tener sólo, se dice, 100 pies de largo. Pero en este supuesto, los cajones archivadores son tremendamente profundos y nos encontraremos con que las partes posteriores de los cajones archivadores alcanzan el pasillo vecino, así… estaremos fuera de las dimensiones en las cuales se pueden alinear los libros. Necesitamos un espacio millón-dimensional para almacenar nítidamente todos los libros, y lo único que tenemos son tres dimensiones: arriba-abajo, derecha-izquierda y delante-detrás. Nuestra pretensión es imaginar un espacio multidimensional en que cada dimensión se disponga en «ángulos rectos» a todas las otras dimensiones. Se puede concebir tal hiperespacio, como se llama, aunque no podamos visualizarlo. Los científicos los utilizan continuamente para organizar la expresión de sus teorías. La geometría de tales espacios —sean o no imaginarios— está siendo bien trabajada y explorada por los matemáticos. En estos espacios lógicos podemos hablar con seguridad sobre localizaciones, caminos, trayectorias, volúmenes (hipervolúmenes), distancias y direcciones.

Ahora ya estamos preparados para ocuparnos de una variación en el tema de Borges, a la que denomino la Biblioteca de Mendel. Esta biblioteca contiene «todos los posibles genomas», o secuencias de ADN. Richard Dawkins ha descrito un espacio similar, al que llama «Tierra biomórfica» en su libro *El relojero ciego*. Su discusión me ha servido de inspiración y tanto la versión de Dawkins como la mía, son del todo compatibles, aunque deseo resaltar algunos puntos que él ha mencionado de pasada.

Si consideramos que la Biblioteca de Mendel está formada por *descripciones* de genomas, entonces podemos compararla con exactitud a la Biblioteca de Babel. El código estándar para describir el ADN consiste en cuatro caracteres, A, C, G y T (iniciales de adenosina, citosína, guanina y timina, los cuatro tipos de nucleótidos que forman las letras del alfabeto del ADN). Todas las permutaciones de 500 páginas de estas cuatro letras están ya en la Biblioteca de Babel. Sin embargo, los genomas típicos son mucho más largos que los libros ordinarios. Aceptando la estimación actual de 3×10^9 nucleótidos en el genoma humano, la descripción exhaustiva de un simple genoma humano —tal como nuestro propio genoma— ocuparía aproximadamente 3.000 de los volúmenes de 500 páginas en la Biblioteca de Babel (manteniendo el mismo tamaño de caja en la impresión)[37]. La descripción del genoma de un caballo (volador o no), de una col o de un pulpo estará compuesta por las mismas letras A, C, G, y T y con seguridad no mucho más largo, así que podemos suponer arbitrariamente que la Biblioteca de Mendel consiste en todas las tiras de ADN descritas en todos las series de 3.000 volúmenes archivados, que están formadas por aquellos cuatro caracteres. De este modo, se conservarán bastantes de los posibles genomas para que sirvan para cualquier propósito teórico serio.

Es cierto que he exagerado la cuestión al describir una Biblioteca de Mendel que contiene «todos los posibles» genomas. Del mismo modo que la Biblioteca de Babel desconocía las lenguas rusa y china, la Biblioteca de Mendel pasa por alto la (aparente) posibilidad de alfabetos genéticos alternativos, basados, por ejemplo, en diferentes componentes químicos. Estamos *aún* empezando por la zona media, asegurándonos de que comprendemos las circunstancias actuales, tanto locales como de la globalidad terrestre, antes de extender nuestras redes más lejos. De este modo, cualquiera de las conclusiones a las que lleguemos referentes a lo que es posible en relación con *esta* Biblioteca de Mendel, deben ser consideradas cuando tratemos de aplicarlas a conceptos de posibilidad más amplios. Esta es realmente una fuerza, más que una debilidad de nuestra táctica, dado que

podemos observar con atención de qué tipo de posibilidad, modesta y circunscrita, estamos exactamente hablando.

Una de las características más importantes del ADN es que todas las permutaciones de las secuencias de adenina, citosina, guanina y timina son también químicamente estables. Todas pueden ser construidas, en principio, en los laboratorios de genética, y una vez construidas, tendrían una vida infinita en sus estantes, como un libro en una biblioteca. Pero no toda secuencia en la Biblioteca de Mendel corresponde a un organismo viable. La mayoría de las secuencias de ADN —la gran mayoría— son seguramente galimatías, unas recetas que no sirven para seres vivos. Esto es lo que Dawkins quiere significar, sin duda, cuando dice que hay muchas más vías de estar muerto (o no vivo) que vías para estar vivo. Pero ¿qué clase de hecho es éste o por qué debe ser así?

3. *La compleja relación entre genoma y organismo*

Si estamos tratando de progresar mediante una atrevida hipersimplificación, debemos estar alerta acerca de algunas de las complicaciones que hemos dejado, por el momento, a un lado. Entiendo que existen tres tipos principales de complejidades que debemos conocer y tener presentes cuando vamos progresando, aunque estemos posponiendo su completa discusión.

La primera concierne a la «lectura» de la «receta». La Biblioteca de Babel presupone la existencia de lectores: la gente que habita en la biblioteca. Sin ellos, la misma idea de la colección de volúmenes no tendría ningún sentido; sus páginas pueden ser embadurnadas con mermelada o algo peor. Si queremos darle algún sentido a la Biblioteca de Mendel debemos presuponer algo análogo a los lectores, porque sin lectores las secuencias de ADN *no especifican* nada; ni ojos azules, ni alas, ni nada por el estilo. Los deconstruccionistas nos dirán que dos lectores de un texto no hacen la misma lectura y esto es cierto cuando consideramos la relación entre el genoma y el ambiente embrionario —el microambiente químico, así como las condiciones de apoyo que le rodean— en el que se ejercen los efectos de la información genómica. El efecto inmediato de la «lectura» del ADN durante la creación de un nuevo organismo es la fabricación de muchas proteínas diferentes a partir de aminoácidos (los cuales deben estar disponibles en la vecindad, naturalmente, listos para juntarse). Hay un número astronómicamente elevado de posibles proteínas, pero aquellas que llegan a ser reales dependen del texto del ADN. Estas proteínas son creadas en estricta secuencia y en cantidades

determinadas por tripletes de nucleótidos, las «palabras», y como ellas, se pueden «leer». Así, en una secuencia de ADN, para especificar lo que se supone que debe especificar, debe existir un elaborado lector-constructor bien provisto de bloques de aminoácidos[38] Pero esto es justamente una pequeña parte del proceso. Una vez que las proteínas han sido creadas, deben ponerse en relación entre sí. El proceso comienza con una sola célula fertilizada que se divide entonces en dos células hijas, que a su vez se vuelven a dividir y así sucesivamente (cada una con su propia copia duplicada de todo el ADN que está siendo leído, naturalmente). Estas recientes nuevas células, de variedades muy diferentes (dependiendo de qué proteínas han sido ubicadas en qué localizaciones y en qué orden), deben a su vez migrar a las localizaciones correctas en el embrión, el cual crece dividiéndose y dividiéndose, construyendo, reconstruyendo, revisando, extendiendo, repitiendo, y así sucesivamente.

Este es un proceso que sólo controla parcialmente el ADN, lo que en efecto *presupone* (y en consecuencia no es en sí mismo *específico*) el lector y el proceso de lectura. Podemos comparar los genomas con una partitura musical. ¿*Especifica* una partitura escrita de la Quinta Sinfonía de Beethoven a esta pieza de música? No lo especificaría para los marcianos, debido a que la partitura presupone la existencia de violines, violas, clarinetes y trompetas. Supongamos que tomamos la partitura y le añadimos una hoja de direcciones y esquemas acerca de cómo construir y tocar todos los instrumentos y la enviamos a Marte. Ahora estamos más cerca de un paquete que podría ser, en principio, utilizado para recrear la música de Beethoven en Marte. Pero los marcianos tendrían que ser capaces de descifrar la receta, hacer los instrumentos y después tocarlos tal como la partitura señala.

Esto es lo que convierte en una fantasía la historia de la novela de Michael Crichton *Parque jurásico* (1990) y la película sobre su argumento realizada por Steven Spielberg: un ADN completamente intacto de un dinosaurio no sería capaz de recrear un dinosaurio sin la ayuda de un lector de ADN de dinosaurio, y aquellos lectores están ya tan extinguidos como los mismos dinosaurios (hay, después de todo, ovarios de dinosaurios). Si *tenemos* un ovario de dinosaurio (vivo), junto con ADN de dinosaurio, puedo *especificar* otro dinosaurio, otro ovario de dinosaurio y así indefinidamente. Pero un ADN de dinosaurio en sí mismo, incluso el ADN completo de un dinosaurio, es solamente la mitad (o dependiendo de cómo se cuente puede ser menos de la mitad) de la ecuación. Podemos decir que toda especie que ya ha existido en este planeta ha tenido su propio dialecto para la lectura del ADN. Aun,

estos dialectos han tenido muchas cosas en común. Los principios de la lectura del ADN son aparentemente uniformes a través de las especies. Esto es lo que hace posible la ingeniería genética, el efecto organísmico de una particular permuta en el ADN puede ser predicho en la práctica. De este modo, la idea de alcanzar, con nuestro propio esfuerzo, un camino que nos lleve a un lector de ADN de dinosaurio es una idea coherente aunque improbable. Con la ayuda de una licencia poética, los realizadores cinematográficos pueden pretender que existe la posibilidad de sustitutos aceptables de los lectores de ADN (introducir el texto del ADN del dinosaurio en el lector del ADN de una rana, y esperar lo mejor)[39].

Nos ayudaremos, también, y con cautela, con alguna licencia poética. Supongamos que procedemos *como si* la Biblioteca de Mendel estuviera equipada con un único lector de ADN, que puede ser un nabo o un tigre, dependiendo de la receta que se encuentre en uno de los volúmenes de genomas. Esto es una hipersimplificación brutal, aunque más tarde podremos reabrir la cuestión del desarrollo y de las complicaciones embrionarias[40]. Cualquiera que sea lector de ADN estándar que escojamos, teniendo en cuenta la astronómica cifra de secuencias de ADN en la Biblioteca de Mendel, será un absoluto galimatías. Cualquier intento de «ejecutar» tal receta para crear un organismo viable terminaría rápidamente en el absurdo. No cambiaríamos este cuadro de manera apreciable si en su lugar nos imaginásemos que hay millones de diferentes dialectos de lectores de ADN, análogos a los diferentes lenguajes reales representados en la Biblioteca de Babel. En esta Biblioteca los libros en inglés pueden ser un galimatías para los lectores polacos y viceversa, pero la inmensa mayoría de los volúmenes son galimatías para todos los lectores. Tomemos un volumen al azar; sin duda podemos imaginar que está compuesto en un lenguaje, el babélico, en el que se relata un cuento maravilloso. (La imaginación es barata si no nos aburrimos con los detalles). Pero si recordamos que las lenguas reales deben ser cosas compactas y *prácticas*, con frases cortas y legibles de las que dependen que se consiga que, con sistemática regularidad, sean transmitidos sus mensajes, podemos estar seguros de que, comparado con la vasta variedad de textos en la biblioteca, las posibles lenguas son extraordinariamente escasas. Así también podemos pretender, por el momento, que cuando hay un lenguaje, hay una especie de lector.

La segunda complejidad que podemos reconocer y posponer se refiere a la viabilidad. Un tigre es viable *ahora*, en ciertas condiciones ambientales existentes en nuestro planeta, pero no habría sido viable en días más

tempranos y puede llegar a ser inviable en el futuro (como de hecho puede ser toda la vida en la Tierra). La viabilidad está relacionada con el medio ambiente en el que el organismo vive. Sin una atmósfera respirable y presas comestibles —ateniéndonos a las condiciones más obvias— las características orgánicas que hacen viables a los tigres hoy serían inútiles. Y dado que el medio ambiente está, en gran parte, compuesto por otros organismos, la viabilidad es una propiedad continuamente cambiante, es decir, una diana móvil, no una condición fija. Este problema se minimiza si siguiendo el ejemplo de Darwin comenzamos en la zona media (de la evolución) con las condiciones ambientales actuales y las extrapolamos cautelosamente hacia posibilidades más tempranas o más tardías. Cabe dejar para más tarde una reflexión acerca del esfuerzo propio inicial que puede (o debe) haber sucedido para poner en marcha esta coevolución de organismos y sus respectivos medios ambientes.

La tercera complejidad concierne a las relaciones entre los textos de los genomas que determinan los organismos viables y las características que estos organismos exhiben. Como hemos señalado varias veces de pasada, no existe un *único* mapa de los nucleótidos, considerados como «palabras» en los genes mendelianos —transportadores putativos de los «specs» (como diría un ingeniero) *para* una determinada característica u otra. No se trata simplemente de que haya una secuencia de nucleótidos para poder deletrear «firme» o «aromático» en el lenguaje del ADN del tomate—, aunque, sin embargo, se puede revisar la secuencia de nucleótidos en este lenguaje para que el efecto sean tomates más firmes y más aromáticos.

Cuando conocemos esta complicación, habitualmente se puntualiza que los genomas no son descripciones o esquemas de productos terminados, sino más bien recetas para construirlos. Esto no significa, como algunos críticos han defendido, que sea siempre —o casi siempre— un error hablar de un gen *para* esto o aquello. La presencia o ausencia de una instrucción en una receta puede marcar una diferencia típica e importante y cualquiera diferencia que se hace puede ser correctamente descrita como *para* lo que la instrucción —es decir, el gen— es. Este punto ha sido con frecuencia tan olvidado por los críticos, que merece una pausa para exponer claramente el error. Richards Dawkins lo ha hecho con un ejemplo tan apropiado que merece la pena ser citado (también resalta Dawkins la importancia de la segunda de nuestras complicaciones, la relación de la viabilidad con el medio ambiente):

La lectura es una habilidad aprendida de una prodigiosa complejidad, pero este hecho no aporta razón alguna para el escepticismo sobre la posible existencia de un gen de la lectura. Todo lo que necesitamos

para establecer la existencia de un gen de la lectura es descubrir un gen para no leer, es decir, un gen que induzca una lesión cerebral que cause una específica dislexia. Una persona disléxica puede ser normal e inteligente en todos los aspectos excepto que no puede leer. Ningún genetista se sorprendería si este tipo de dislexia se transmitiese de acuerdo con las leyes mendelianas. Es obvio que, en esta situación, el gen solamente exhibiría sus efectos en un medio ambiente que incluyese una educación normal. En un medio ambiente prehistórico no provocaría efectos detectables o quizás algunos efectos diferentes que hubieran sido conocidos para los genetistas habitantes de las cavernas, por ejemplo, un gen que incapacite la lectura de las huellas de los animales. En nuestro cultivado ambiente sería apropiadamente denominado gen de la dislexia, dado que ésta sería su consecuencia más llamativa. De modo similar, un gen que causa una ceguera total también impediría la lectura, pero no sería de utilidad considerarlo como un gen para no leer. Esto es simplemente debido a que prevenir leer no sería su más obvio o debilitante efecto fenotípico (Dawkins 1982:23. Véanse también Dawkins 1989a:281-282, y Sterelny y Kitcher 1988).

La vía indirecta en la que grupos de codones —que son tripletas de nucleótidos del ADN— instruyen el proceso constructivo no nos prohíbe, por lo tanto, hablar de un gen para *x* o para *y*, utilizando el lenguaje taquigráfico familiar de los genetistas y teniendo presente qué es lo que estamos haciendo. Pero esto significa que hay diferencias fundamentales entre el espacio de los genomas y el espacio de los «posibles» organismos. El hecho de que *nosotros* podamos describir un producto terminado —es decir, una jirafa con rayas verdes en lugar de manchas marrones— no garantiza que haya una receta de ADN para hacerla. Esto puede ser así porque, debido a los peculiares requerimientos del desarrollo, no hay un punto de partida en el ADN que tenga como destino una jirafa de este tipo.

Esto no es nada convincente. ¿Qué hay de imposible en una jirafa con rayas verdes? ¡Las cebras tienen rayas, los ánades tiene plumas verdes en las cabezas! No hay nada biológicamente imposible acerca de las propiedades aisladas, y seguramente también podrían aplicarse conjuntamente a una jirafa. Podríamos razonar así, pero no es posible. También podemos pensar que es posible un animal con rayas y una cola con manchas, pero lo más seguro es que no exista. James Murray [1989] desarrolló modelos matemáticos que muestran cómo el proceso de desarrollo en la distribución de colores en los animales puede fácilmente terminar con un animal con manchas y una cola con rayas, pero no viceversa. Este dato es sugestivo, aunque aún no es extrictamente una prueba de imposibilidad. Cualquiera que haya aprendido a construir un barquito en el interior de una botella —una habilidad bastante difícil— puede pensar que es del todo imposible colocar una pera fresca en el interior de una botella de cuello estrecho, aunque no es así; el ejemplo son las botellas del licor *Poire William*. ¿Cómo se hace? ¿Puede el vidrio caliente ser soplado alrededor de la pera sin quemarla? No, las botellas se cuelgan en los árboles en primavera de modo que las peras crecen dentro de ellas. Probar que

no hay vía biológica directa para resolver algún problema no es nunca prueba de imposibilidad. ¡Recordemos la segunda regla de Orgel!

En su descripción del territorio biomórfico Dawkins subraya que un mínimo cambio en el genotipo (la receta) puede producir un llamativo y amplio cambio en el fenotipo (el organismo individual resultante), aunque tiende a minusvalorar las mayores implicaciones de ese hecho: si un simple paso en el genotipo puede producir un paso gigante en el fenotipo, los pasos intermedios para el fenotipo pueden estar simplemente no disponibles, dadas las reglas del «mapping» genómico. Utilicemos un ejemplo deliberadamente extremado y divertido, pensemos que si un animal puede tener colmillos de veinte y de cuarenta centímetros, también podría tener colmillos de treinta centímetros, pero las reglas para el desarrollo de colmillos en el sistema de recetas pueden no permitir tal eventualidad. La especie en cuestión puede haber tenido que escoger entre colmillos «demasiado cortos» por diez centímetros o «demasiado largos» en diez centímetros. Esto significa que al utilizar postulados de la ingeniería sobre el diseño óptimo, se debe ser muy cauteloso al asumir que lo que parece intuitivamente disponible o posible es en realidad accesible en el espacio de diseño de los organismos, teniendo en cuenta cómo se leen estas recetas. (Este será un tema principal en los capítulos 8, 9 y 10).

4. *La posibilidad hecha naturaleza*

Con la ayuda de la Biblioteca de Mendel podernos resolver ahora —o al menos unificar bajo una perspectiva singular— algunos de los problemas más insistentes de las «leyes biológicas» y de lo que es posible, imposible y necesario en el mundo. Recordemos la necesidad de aclarar estos temas, porque si vamos a explicar la vía por la que las cosas *son*, debe de ser en un trasfondo de cómo las cosas *pueden* haber sido, o *deben* ser o *no podrían ser*. Es el momento de definir un concepto restringido de posibilidad biológica: «*x* es biológicamente posible si *y* sólo si *x* es una ejemplificación de un genoma accesible o una característica de sus productos fenotípicos».

¿Accesible desde dónde? ¿Mediante qué procesos? ¡He aquí la cuestión! Tenemos que especificar un punto de partida en la Biblioteca de Mendel y un medio de «viaje». Supongamos que vamos a comenzar a partir de hoy. Entonces hablaremos primero acerca de lo que es posible *ahora*; es decir, en el futuro próximo, utilizando cualquier medio de viaje disponible en la actualidad. Contamos como posible todas las especies contemporáneas *reales*

y todas sus características —incluyendo las características que poseen en virtud de sus relaciones con otras especies y sus características—, más lo que se pueda obtener viajando en el «curso de la naturaleza» —sin manipulación humana—, o bien con la ayuda de grúas artificiales como las técnicas tradicionales de crías de animales (y en este tema, de la cirugía) o a través de los innovadores vehículos de la ingeniería genética. Después de todo, nosotros, seres humanos, y todas nuestras habilidades somos productos de la biosfera contemporánea. De este modo es biológicamente posible comerse un pavo fresco en la comida de Navidad del año 2001 sí, y sólo sí, al menos un genoma de pavo ejemplificado ha producido los efectos fenotípicos requeridos a tiempo para dicha comida. Es biológicamente posible para nosotros montarnos en un «pteranodon» antes de morir si y sólo si la tecnología del tipo Parque Jurásico permite que *este* tipo de genoma pueda expresarse a tiempo.

No importa cómo agrupemos estos parámetros de «viaje», la noción resultante de posibilidad biológica tendrá una importante propiedad: algunas cosas serán «más posibles» que otras; esto es, más cercanas en el espacio multidimensional de la búsqueda, y más accesibles, «más fáciles» de conseguir. Las cosas que habrían sido vistas como imposibilidades biológicas hace pocos años —como son las plantas que crecen en la oscuridad en virtud de poseer genes de luciérnagas— no sólo son posibles ahora, sino reales. ¿Son posibles dinosaurios en el siglo XXI? Bien, los vehículos *para llegar allí* desde aquí han sido desarrollados hasta un punto desde donde podemos al menos contar una historia extraordinariamente buena; una historia que requiere una extraordinaria y pequeña licencia poética. («Allí» es una parte de la Biblioteca de Mendel a través de la cual el árbol de la vida detuvo su extensión hace sesenta millones de años).

¿Qué reglas gobiernan el viaje a través de este espacio? ¿Qué reglas o leyes dificultan las relaciones entre los genomas y sus productos fenotípicos? Hasta ahora, todo lo que hemos conocido son necesidades lógicas y matemáticas por un lado y leyes de la física por otro. Esto es, hemos procedido como si conociéramos cuáles son las posibilidades lógicas y (meramente) físicas. Estos son temas difíciles y controvertidos, aunque podemos considerarlos *afianzados*: simplemente asumimos alguna versión establecida de aquellas variedades de posibilidad y necesidad, y entonces desarrollamos nuestra noción de posibilidad biológica restringida en estos términos. La ley de los grandes números y la ley de la gravedad, por ejemplo, están consideradas como asentadas sin reserva y eternamente en el espacio. El

afianzamiento de las leyes físicas nos lleva a decir simple y llanamente, por ejemplo, que los diversos genomas son físicamente posibles; porque la química nos dice que todos, si se encuentran, son estables.

Manteniendo la lógica, la física y la química afianzadas conjuntamente, podemos escoger un punto de partida diferente. Podemos optar por algún momento en la Tierra, hace quinientos mil millones de años, y considerar lo que era entonces biológicamente posible. No demasiado, porque antes de que los tigres llegaran a ser posibles (en la Tierra), los eucariotas, y las plantas que producen el oxígeno atmosférico en grandes cantidades, y muchas otras cosas, tenían que llegar a ser reales. En una visión retrospectiva podemos decir que los tigres fueron, de hecho, posibles desde el principio, aunque fuera a distancia y extremadamente improbables. Una de las virtudes de este modo de pensar acerca de la posibilidad es que une sus fuerzas con la probabilidad, permitiéndonos negociar con claridad las declaraciones del tipo todo o nada acerca de la posibilidad de afirmaciones respecto a distancias relativas, que es lo que nos interesa para la mayoría de los propósitos. (Las declaraciones todo o nada acerca de la posibilidad biológica eran casi imposibles —¡vaya, de nuevo la palabra!— de adjudicar, así que esto no se ha perdido). Como hemos visto en nuestra exploración de la Biblioteca de Babel, no es muy diferente nuestro veredicto acerca de si es «posible en principio» encontrar algún libro concreto en ese vasto espacio. Lo que importa es que es prácticamente posible, en uno u otro sentido del adjetivo «práctico»; que cada uno escoja el que le guste.

En realidad, esta no es ciertamente una definición estándar de posibilidad, ni incluso un *tipo* estándar de definición de posibilidad. La idea de que algunas cosas pueden ser «más posibles» que otras (o más posibles desde aquí que desde allí) está en desacuerdo con una comprensión estándar del término, y algunas críticas filosóficas pueden decir que ésta no es simplemente una definición de *posibilidad*, cualquiera que sea. Otros filósofos han defendido el concepto de posibilidad comparativa (véase especialmente Lewis 1986:IOSS.), aunque no deseo combatir esta idea. Si esto no es una descripción de posibilidad, que así sea. Esta es una propuesta de *sustitución* de una definición de posibilidad. Quizá, después de todo, no necesitamos el concepto de posibilidad biológica (con su requerida aplicación todo o nada) para cualquier propósito de investigación seria. Quizá todo lo que necesitamos es el grado de accesibilidad en el espacio de la Biblioteca de Mendel y, de hecho, es un mejor concepto de posibilidad de lo que sería cualquiera versión todo o nada. Estaría bien, por ejemplo, obtener alguna forma de clasificar por

grados los siguientes ejemplos de acuerdo con su posibilidad biológica: tomates de diez libras, perros acuáticos, caballos voladores, árboles voladores.

Esto no sería suficiente para satisfacer a muchos filósofos y sus objeciones serían serias. Si las consideramos brevemente, al menos, veremos con mayor claridad lo que defiendo y lo que no defiendo. En primer lugar, ¿no es entrar en un círculo vicioso tratar de definir la posibilidad en términos de accesibilidad? ¿No reintroduce el último término al primero en su sufijo, mientras que aún permanece indefinido? Bueno, no del todo. Deja alguna cuestión definitivamente indefinida, cosa que ya sabíamos antes de empezar. Hemos supuesto que poseíamos uno u otro concepto de posibilidad *física*, afianzada en el momento actual: nuestra idea de accesibilidad presupone que esta posibilidad física, cualquiera que sea, nos deja libertad de acción; algunas vías abiertas (no sólo una vía) en el espacio. En otras palabras, aceptamos la presunción de que nada nos detendrá para seguir adelante en las vías que están abiertas relativas a la física**[41]**.

Los interrogantes de Quine (en la cabecera de este capítulo) nos inducen a preocuparnos acerca de si podríamos contar con objetos posibles y no reales. Una de las virtudes del tratamiento propuesto para la posibilidad biológica es que, gracias a su «arbitrario» sistema formal —el sistema que nos ha sido propuesto arbitrariamente e impuesto por la naturaleza al menos en nuestro «rincón del bosque»—, podemos contar los diferentes posibles genomas no reales; son muchísimos aunque finitos en número y no hay dos que sean exactamente iguales. (Por definición, los genomas son distintos si no comparten un nucleótido en uno de los varios miles de millones de *loci*,) ¿En qué sentido son *realmente* posibles los genomas no reales? Tan sólo en este sentido: *si* llegaran a formarse, serían estables. Pero si por una conspiración de acontecimientos pudieran o no pudieran llegar a ser formado los genomas, esta es otra cuestión, que debe ser tratada en términos de accesibilidad desde una u otra localización. Podemos asegurar que la mayoría de los genomas en este grupo de posibilidades estables *nunca* serán formados, dado que la muerte por el calor del universo sorprenderá el proceso de construcción antes de que haya conseguido un considerable hueco en el espacio.

Otras dos objeciones a esta propuesta sobre posibilidad biológica han sido expuestas. En primer lugar, ¿no es una propuesta ultrajante «centrada en los genes», hacer depender *todas* las consideraciones de la posibilidad biológica de la accesibilidad a uno u otro genoma en la Biblioteca de Mendel? El tratamiento que nosotros proponemos para la posibilidad biológica ignora de plano (y en consecuencia establece implícitamente como imposible) a las

«criaturas» que no sean puntos finales de alguna rama de ese árbol de la vida que ya nos ha llevado tan lejos, al punto en que nos encontramos hoy. ¡Esta *es* precisamente la gran unificación de la biología que descubrió Darwin! A menos que alberguemos fantasías acerca de la creación espontánea de nuevas formas de vida por un «creador especial» o, según la secular versión de los filósofos, por «coincidencia cósmica», aceptamos que toda característica de la biosfera es un fruto u otro del árbol de la vida (o, si no de nuestro propio árbol de la vida, de algún otro árbol de la vida, con sus propias relaciones de accesibilidad). Ningún hombre es una isla, proclamó John Donne, y Charles Darwin añadió que tampoco lo es una almeja o un tulipán; todo ser vivo posible está conectado mediante istmos de descenso con el resto de seres vivos. Obsérvese que esta doctrina rige para cualquier maravilla que la tecnología pueda producir en el futuro, a condición de que —como ya hemos señalado— los propios técnicos, sus herramientas y sus métodos, estén firmemente localizados en el árbol de la vida. Determinar la existencia de vida en formas del espacio exterior es un pequeño paso posterior, a condición de que, también, sean los productos de un árbol de la vida enraizado, como nosotros, en algún fundamento físico no milagroso. (Este tema será analizado en el capítulo 7.)

En segundo lugar, ¿por qué debemos tratar de manera tan diferente la posibilidad biológica de la posibilidad física? Si asumimos que las «leyes de la física» fijan los límites de la posibilidad física, ¿por qué no intentar definir la posibilidad biológica de acuerdo con «leyes de la biología»? (Volveremos a examinar las leyes físicas y la necesidad física en el capítulo 7, mientras tanto la diferencia parece grande). Muchos biólogos y filósofos de la ciencia han sostenido que hay leyes biológicas. ¿No las elimina la definición propuesta? O bien, ¿no las declara superfluas? No son eliminadas. Se permite a alguien argumentar a favor del dominio de alguna ley de la biología sobre el espacio de la Biblioteca de Mendel, aunque aplica un difícil peso de prueba sobre cualquiera que sostenga que *por encima* de la leyes de las matemáticas y de la física hay leyes de la biología. Consideremos, por ejemplo, la suerte de la ley de Dollo:

La «ley de Dollo» establece que la evolución es irreversible… (Pero) no hay razón alguna por la que las tendencias generales en la evolución no puedan ser reversibles. Si hay una tendencia hacia las grandes cornamentas durante la evolución, puede haber fácilmente de nuevo una posterior tendencia hacia cornamentas más pequeñas. La «ley de Dollo» es en realidad una declaración acerca de la improbabilidad estadística de seguir exactamente la misma trayectoria evolucionaría dos veces (o cualquier trayectoria *particular*) en ambas direcciones. Un simple paso mutacional puede ser fácilmente revertido. Pero para cantidades mayores de pasos mutacionales… el espacio matemático de todas las posibles trayectorias es tan vasto que la probabilidad de que dos trayectorias lleguen alguna vez al

mismo punto llega a ser «evanescentemente» pequeña… No hay nada misterioso o místico en la «ley de Dollo», ni es algo que podamos ir y «probarlo» en la naturaleza. Se deriva simplemente de las leyes elementales de la probabilidad (Dawkins 1986:94).

Hay muchos candidatos que defienden el papel de «ley biológica irreductible». Por ejemplo, muchos han argumentado que hay «leyes del desarrollo» o «leyes de la forma» que dificultan la relación entre genotipo y fenotipo. A su debido tiempo consideraremos sus condiciones, pero ya podemos localizar al menos algunas de las dificultades más llamativas sobre la posibilidad biológica, no como «leyes de la biología» sino como características ineludibles de la geometría del espacio de diseño, como la ley de Dollo (o la ley de Hardy-Weinberg sobre la frecuencia de los genes, que es otra aplicación de la teoría de la probabilidad, pura y simplemente).

Examinemos el caso de las aves cornudas. Como ha señalado Maynard Smith, no hay ninguna y no sabemos por qué. ¿Puede ser a causa de que han sido descartadas por una *ley* biológica? ¿Son totalmente imposibles las aves cornudas? ¿Deben ser inviables en alguno o en todos los posibles ambientes, o es que no existe una vía «desde aquí hasta allí» debido a las restricciones en el proceso de la lectura del genoma? Como ya hemos señalado, las graves restricciones para este proceso deben impresionarnos, pero no debemos perder la calma. Aquellas restricciones pueden no ser una característica *universal* sino una característica espacialmente local, análoga a lo que Seymour Papert ha denominado como el fenómeno QWERTY en la cultura de los ordenadores y de los teclados.

En la línea superior de las letras de un teclado estándar se lee QWERTY. Para mí esto simboliza la vía por la que la tecnología puede muy a menudo servir no como una fuerza de progreso sino para mantener la confusión en las cosas. La disposición QWERTY no tiene una explicación racional, sólo histórica. Fue introducida como respuesta a un problema en los inicios de la máquina de escribir. Las teclas solían atascarse. La idea fue minimizar el problema de la colisión separando aquellas teclas que seguían una u otra frecuencia… Una vez adoptada, el resultado fue su aplicación en muchos millones de máquinas de escribir… el coste social del cambio… aumentó con el interés creado por el hecho de que muchos dedos conocían ahora el teclado QWERTY. QWERTY sigue existiendo a pesar de la existencia de otros sistemas más «racionales» (Papert 1980:33)[42].

Las imperiosas restricciones que encontramos en la Biblioteca de Mendel pueden parecer leyes universales de la naturaleza desde nuestra miope perspectiva, pero desde una perspectiva diferente pueden aparecer como condiciones locales con explicaciones históricas[43]. Por eso, el tipo de definición que deseamos es el de un concepto restringido de la posibilidad biológica; el ideal de un concepto universal de la posibilidad biológica sería interpretado equivocadamente. Pero como ya he admitido, esta posición no

descarta las leyes biológicas; simplemente coloca el peso de la prueba sobre aquellos que desean proponer alguna. Y mientras tanto, nos facilita una estructura para describir sobre ella amplias e importantes clases de regularidad que descubrimos en los modelos de *nuestra* biosfera.

CAPÍTULO 6

Líneas de realidad en el espacio de diseño

I. *A la deriva por el espacio de diseño*

> Los animales reales que alguna vez han vivido en la Tierra son una muy pequeña parte de los animales teóricos que *podrían haber* existido. Estos animales reales son los productos de un número muy pequeño de trayectorias evolutivas a través del espacio genético. La inmensa mayoría de las trayectorias teóricas a través del espacio animal daría origen a monstruos imposibles. Los animales reales se distribuyen aquí y allá, entre los hipotéticos monstruos, cada uno ubicado en su propio y único lugar en el hiperespacio genético. Cada animal real está rodeado por un pequeño grupo de vecinos, la mayoría de los cuales nunca ha existido, aunque unos pocos tienen sus ancestros, sus descendientes y sus primos.
>
> Richard Dawkins, *El relojero ciego*

Los genomas reales que han existido alguna vez son un subgrupo extraordinariamente pequeño de todos los genomas combinatoriamente posibles, del mismo modo que los libros reales en las bibliotecas del mundo son un pequeño subgrupo evanescente de los libros de la imaginaria Biblioteca de Babel. Cuando revisamos la Biblioteca de Babel nos sorprende lo difícil que resulta especificar una *categoría* de libros que, aunque no es muy numerosa, es extraordinariamente pequeña si se la compara con la totalidad. El grupo de libros compuesto en su totalidad por frases gramaticales inglesas es un *vasto* pero evanescente subgrupo y el grupo de los libros legibles que tienen sentido es un *vasto* aunque evanescente subgrupo de aquel. Oculto en este subgrupo se encuentra el vasto grupo de libros sobre la gente cuyo nombre es Charles, y dentro de este grupo (aunque *evanescentemente* difícil de encontrar) está el *vasto* grupo de libros que comparten el propósito de decir la verdad acerca de Charles Darwin, y un vasto pero evanescente subgrupo de éstos está formado por libros escritos en verso. Y así, sucesivamente. El número de libros reales dedicados a Charles Darwin es muy elevado, pero no es un número vasto y no conseguiremos este grupo (este grupo tal como es hoy o como será en el año 3000 d. C.) intensificando de este modo los adjetivos restrictivos. Para llegar hasta los

libros reales tenemos que retroceder hasta el proceso histórico que los ha creado, con toda su desaliñada particularidad. Esto mismo es aplicable, en realidad, a los organismos reales o sus genomas reales.

No necesitamos leyes biológicas para impedir que la mayoría de las posibilidades físicas lleguen a ser reales; para la mayoría ha sido así por una evidente ausencia de oportunidad. La única «razón» por la que *todos* nuestros tías y tíos no reales nunca llegaron a existir, ha sido que nuestros abuelos no tuvieron tiempo o energía (para no hablar de la inclinación) para crear más que unos pocos genomas vecinos. Entre los muchos seres no reales posibles, algunos son —o eran— «más posibles» que otros: es decir, su apariencia era más *probable* que la apariencia de otros, simplemente porque eran *vecinos* de genomas reales, situados sólo a unas cuantas opciones de distancia en el proceso casual que, como un cierre de cremallera, conjuga un nuevo volumen de ADN a partir de los anteproyectos de los padres, o sólo uno o unos pocos errores tipográficos hechos al azar en el gran proceso de la copia. ¿Por qué no suceden los cuasi-errores? No hay razón; precisamente no sucede que sucedan. Y entonces, cuando los genomas reales que *sucede* que sucedan comienzan a moverse desde sus localizaciones en el espacio de diseño de los cuasi-errores, su probabilidad de suceder se hace aún más pequeña. ¡Estuvieron muy cerca de llegar a ser reales pero su momento pasó! ¿Tendrán otra oportunidad? Es posible pero altamente improbable, dado el vasto tamaño del espacio en el que residen.

Pero ¿qué fuerzas, si hay alguna, apartan los caminos de la realidad cada vez más lejos de sus localizaciones? El movimiento que ocurre, en ausencia de fuerza alguna, se denomina *arrastre genético aleatorio*. Se puede pensar que este arrastre genético, al ser aleatorio, tendería siempre a finalizar por sí mismo, volviendo a los mismos genomas una y otra vez en ausencia de fuerzas selectivas, pero el propio hecho de que haya sólo un muestreo limitado en el enorme espacio (el cual —¡recuérdese!— tiene un millón de dimensiones) conduce inevitablemente a la acumulación de «distancia» entre los genomas reales (la conclusión de la ley de Dollo).

La tesis central de Darwin es que cuando la fuerza de la selección natural se impone en esta corriente aleatoria que serpentea formando meandros, al arrastre genético por la corriente se suma la elevación. Cualquier movimiento en el espacio de diseño puede ser medido, pero el movimiento de arrastre aleatorio es, intuitivamente, movimiento a uno y otro lado que no consigue nada importante. Considerado como un trabajo de I+D es lento y *perezoso* y conduce a la acumulación de simples *cambios tipográficos*, pero no a la

acumulación de *diseños*. De hecho, es peor que esto, ya que por más que la mayoría de las mutaciones —errores— sean neutrales, la mayoría de los errores que no son neutrales tendrán efectos deletéreos. En ausencia de selección natural, el arrastre (genético aleatorio) en el espacio de diseño se hace inexorablemente *cuesta abajo*. La situación en la Biblioteca de Mendel es igual que la situación en la Biblioteca de Babel. La mayoría de los cambios tipográficos en *Moby Dick* pueden considerarse en la práctica como neutrales, es decir, tan buenos como invisibles para la mayoría de los lectores; de los pocos cambios que marcan algunas diferencias, la mayor parte *dañará* el texto, empeorándolo y convirtiéndolo en un relato menos coherente y menos comprensible. Sin embargo, consideremos como un ejercicio la versión del juego de Peter De Vries en la que el objetivo es *mejorar* un texto mediante un simple cambio tipográfico. No es imposible, aunque está muy lejos de ser fácil.

Estas intuiciones acerca de conseguir algún lugar importante, en relación con la *mejoría* del diseño y con la *elevación* en el espacio de diseño, son poderosas y familiares, pero ¿son dignas de confianza? ¿Se trata, quizá, de un confuso legado de la visión predarwiniana del diseño que llega de un Dios Hacedor? ¿Cuál es la relación entre las ideas de diseño y progreso? Entre los teóricos evolucionistas no existe un acuerdo definitivo en este punto. Algunos biólogos son puntillosos, por lo que en sus discursos se desvían por largos caminos con el fin de evitar las alusiones al diseño o función en su propio trabajo, mientras que otros construyen la totalidad de su carrera alrededor del análisis funcional de esto o de esto otro (un órgano, estructuras de huellas, «estrategias» reproductivas, etc.). Otros biólogos piensan que podemos hablar de diseño o de función sin comprometernos con ninguna dudosa doctrina sobre el progreso. Otros no están tan seguros. ¿Asestó Darwin «un golpe mortal a la teleología», como Marx exclamó, o demostró que el «significado racional» de las ciencias naturales se explica empíricamente (Marx tenía razón) construyendo en la ciencia una casa segura para la discusión funcional o teleológica?

¿Es el diseño algo que puede ser medido aunque sea de manera indirecta e imperfecta? Es bastante curioso que el escepticismo sobre esta perspectiva socave realmente la más potente fuente de escepticismo ante el darwinismo. Como he subrayado en el capítulo 3, los retos más poderosos para el darwinismo siempre han adoptado esta forma de expresión: ¿son los mecanismos darwinianos lo bastante poderosos, o lo bastante eficientes, como para haber hecho todo este trabajo en el tiempo disponible? ¿Qué trabajo? Si

la pregunta se refiere al mero arrastre lateral en el espacio tipográfico de los posibles genomas, la respuesta será obvia e incontrovertible: sí, ha habido bastante *más* que bastante tiempo. La velocidad con la que el arrastre genético aleatorio debe acumular distancia tipográfica puede ser calculada, dándonos una suerte de velocidad límite, y tanto la teoría como la observación coinciden en que la evolución real sucede más lentamente que ésta[44]. Los «productos» que impresionan a los escépticos no son las diferentes cadenas de ADN en sí mismas, sino los organismos admirablemente intrincados, complejos y bien diseñados cuyos genomas son esas cadenas.

Ningún análisis de los genomas aislados de los organismos que han creado puede proporcionarnos la dimensión que estamos buscando. Sería como tratar de definir la diferencia entre una buena novela y una gran novela según las frecuencias relativas de los caracteres alfabéticos que en ellas se encontraran. El organismo debe contemplarse en su totalidad, en su medio ambiente, para que podamos conseguir algún beneficio en este asunto. Tal como pensaba William Paley, lo que es verdaderamente impresionante es la abundancia asombrosa de ingenio y dispositivos que funcionan como la seda en la materia que compone los seres vivos. Y cuando volvemos a examinar el organismo, nos encontramos de nuevo con que la mera tabulación de las partes que lo componen no nos proporciona lo que deseamos.

¿Cuál puede ser la relación entre la cantidad de complejidad y la cantidad de diseño? «Menos es más», dijo el arquitecto Mies van der Rohe. Recordemos el caso del famoso motor fuera borda británico Seagull, un triunfo de la sencillez, un diseño que honra el principio que dice que «lo que no se encuentra allí no puede romperse». Nos gustaría ser capaces de conocer —e incluso medir, si fuera posible— la excelencia del diseño puesta de manifiesto en el tipo apropiado de simplicidad. Pero ¿cuál es el tipo apropiado? O, dicho de otro modo, ¿cuál es el tipo de *ocasión* apropiada para la simplicidad? No cualquier ocasión. Algunas veces más es más, y naturalmente lo que hace que el motor británico Seagull sea tan admirable, es la elegante conjunción entre complejidad y simplicidad; nadie tiene una estimación tan alta por una hélice.

Podemos empezar a conseguir una visión clara de esta cuestión si reflexionamos acerca de la evolución convergente y de las ocasiones en las que esto ocurre. Y, como es algo tan frecuente, el escoger ejemplos extremos —e imaginarios— es un buen camino para dedicar nuestra atención a lo que importa. En este ejemplo, un caso extremo favorable para ser considerado es la vida extraterrestre y, naturalmente, ésta puede algún día, quizá pronto,

pasar de fantasía a realidad, si el programa conocido como SETI («Search for Extra-Terrestrial Intelligence») encuentra algo. Si la vida en la Tierra es sólidamente contingente —si su mera ocurrencia en cualquier forma es un feliz accidente—, entonces, ¿qué es lo que podemos decir, si podemos decir algo, acerca de la vida en otros planetas del universo? Podemos establecer *algunas* condiciones, con una confianza próxima a la certeza. Estas condiciones, al principio, parecen corresponder a dos grupos bien contrastados: necesidades y lo que podemos llamar «situaciones óptimas».

Establezcamos en primer lugar una necesidad. La vida, en cualquier lugar, consiste en entidades que poseen un metabolismo autónomo. Algunos dicen que esto es «verdad por definición». Definiendo la vida de este modo, se excluye a los virus como formas vivas, mientras que las bacterias se mantienen dentro del atractivo círculo. Pueden existir buenas razones para definir la vida de este modo, aunque creo que comprenderíamos más claramente la importancia del metabolismo autónomo si lo entendiésemos como una profunda, si no absolutamente necesaria, condición para el tipo de complejidad que es necesaria para detener los efectos demoledores de la segunda ley de la termodinámica. Todas las complejas estructuras macromoleculares tienden a romperse con el tiempo, así que, a menos que un sistema sea un sistema *abierto*, capaz de tomar materiales frescos y reabastecerse por sí mismo, su carrera será corta. La pregunta ¿qué hace vivir?, puede tener diferentes respuestas en diferentes planetas, aunque no traiciona una hipótesis «geocéntrica», o si se quiere «antropomórfica».

¿Qué diremos sobre la visión? Sabemos que los ojos han evolucionado muchas veces de modo independiente, aunque la visión no es una necesidad en la Tierra, dado que las plantas viven perfectamente sin visión. Sin embargo, un poderoso argumento puede presentarse diciendo que si un organismo va a llevar adelante sus proyectos metabólicos movilizándose, y si el medio en el que la locomoción tiene lugar es transparente o translúcido y está muy iluminado por la luz ambiental, en estas condiciones, dado que la locomoción funciona bastante mejor (con respecto a objetivos de autoprotección, metabólicos y reproductivos) si el que se mueve es guiado gracias a la información que recibe de los objetos distantes, y considerando que tal información puede ser acumulada por la visión de alta fidelidad y a bajo coste, la visión es una buena apuesta. Así, no nos sorprenderá si se comprueba que los organismos locomotores de otros planetas (con atmósfera transparente) tienen ojos. Los ojos son, obviamente, una gran solución para un problema muy general que pueden encontrarse con frecuencia los seres

con metabolismo dotados de movilidad. Los ojos no siempre están «disponibles», naturalmente, por razones QWERTY, pero son soluciones racionales para este problema de diseño extraordinariamente abstracto.

2. *Movimientos forzados en el juego del diseño*

Ahora que hemos sentido atracción por lo que es obviamente racional, cuando coincide un conjunto general de circunstancias, podemos reflexionar y comprobar que nuestro argumento a favor de la necesidad de poseer un metabolismo autónomo puede ser considerado como la *única* solución aceptable para el problema *más general* en el diseño de la vida. Si queremos vivir, tenemos que comer. En el ajedrez, cuando sólo existe una jugada para evitar el desastre, a ésta se la denomina un *movimiento forzado*. Pero tal movimiento no viene forzado por las reglas del ajedrez y tampoco, ciertamente, por las leyes de la física (siempre estamos a tiempo de darle un puntapié al tablero y salir corriendo), sino por lo que Hume llamó un «dictado de la razón». Es totalmente obvio que hay una, y sólo una, solución, como cualquier persona con dos dedos de frente puede fácilmente comprobar. Otras alternativas conducirían al suicidio inmediato.

Además de poseer un metabolismo autónomo, cualquier organismo debe tener también unas fronteras más o menos definidas, que le distingan de cualquier otra entidad. Esta exigencia tiene, además, una obvia y apremiante razón de ser: «Tan pronto como algo se encuentra implicado en el problema de la autopreservación, las fronteras se convierten en importantes, ya que, si bien estamos obligados a preservarnos a nosotros mismos, no deseamos malgastar esfuerzos tratando de preservar la totalidad del mundo: por eso es por lo que trazamos la línea» (Dennett 1991a:174). También debemos esperar que los organismos dotados de locomoción que vivan en otro planeta tengan sus límites eficientemente trazados, como los tienen los organismos terrestres. ¿Por qué? (¿Por qué en la Tierra?). Si no se tuviera en cuenta el coste, no se consideraría el perfil hidrodinámico de los organismos que se mueven en un fluido relativamente denso, como el agua. Pero el coste es *siempre* un objeto: la segunda ley de la termodinámica lo garantiza.

De este modo algunas «necesidades biológicas», al menos, pueden ser consideradas como soluciones obvias para los problemas más generales, es decir, como *movimientos forzados en el tablero del espacio de diseño*. Hay casos en los que, por una u otra razón, sólo existe una vía por la que se pueden hacer las cosas. Pero las razones pueden ser profundas o superficiales.

Las razones profundas son las limitaciones impuestas por las leyes de la física, tal como sucede en la segunda ley de la termodinámica o por las leyes matemáticas y lógicas[45]. Las razones superficiales son precisamente razones históricas. Suele haber una o dos vías por las que este problema puede solucionarse, pero ahora que algunos antiguos accidentes históricos nos han descartado una de las vías, sólo la otra está remotamente disponible; ha llegado a ser una «necesidad virtual», una necesidad para todos los propósitos prácticos, teniendo en cuenta las cartas que se han jugado. En realidad, las otras opciones han dejado de serlo.

El matrimonio entre el azar y la necesidad es la marca característica de las regulaciones biológicas. A la gente, a menudo, le gustaría preguntar: «¿Es simplemente un hecho contingente sólido que las circunstancias sean como son, o podemos leer alguna necesidad profunda en ellas?». La respuesta casi siempre es ésta: ambas cosas. Pero el tipo de necesidad que se acomoda tan bien con el azar de la ciega generación aleatoria es la necesidad de *razón*. Esta es una ineludible variedad teleológica de necesidad, el dictado de lo que Aristóteles llamó *razonamiento práctico* y Kant denominó *imperativo hipotético*: «Si *deseas conseguir el objetivo G*, entonces esto es lo que *debes* hacer, dadas las circunstancias».

Mientras más universales son las circunstancias, más universal es la necesidad. Por eso, no debe sorprendernos encontrar que los seres vivos en otros planetas incluyan un sistema locomotor y estén dotados de ojos, y aún nos sorprendería más —nos dejará sin habla— si nos encontrásemos con algo que corriese precipitadamente a nuestro alrededor con varios diseños, pero desprovisto de procesos metabólicos. Analicemos ahora la diferencia entre las similitudes que nos sorprenderían y las similitudes que no lo harían. Supongamos que el ya citado programa SETI consigue abundantes fondos y llega a establecer comunicaciones con seres inteligentes en otros planetas. No nos causaría sorpresa que aquellos seres comprendiesen y utilizasen la misma aritmética que nosotros. ¿Por qué no? Porque la aritmética es *correcta*.

¿No pueden existir diferentes clases de sistemas aritméticos, todos igualmente buenos? Marvin Minsky, uno de los fundadores de la inteligencia artificial, ha analizado esta curiosa cuestión y su respuesta ingeniosamente razonada es no. En «Why Inteligent Aliens Will Be Intelligible?». (‘¿Por qué los alienígenas inteligentes serán inteligibles?’), Minsky ofrece fundamentos para creer en lo que él llama el

principio de la escasez: ¡cada vez que dos procesos relativamente simples tienen productos que son similares, es probable que aquellos productos sean completamente idénticos! (Minsky 1985a:119).

Consideremos el conjunto de todos *los procesos posibles*, los cuales Minsky interpreta *a la* Biblioteca de Babel como todas las permutaciones de todos los posibles ordenadores. (Cualquier ordenador puede ser identificado, de modo abstracto, como una «máquina de Turing» u otro modelo, y a estos ordenadores les puede adjudicar un único número para su identificación, y, así, serán colocados en orden numérico de modo similar al orden alfabético en la Biblioteca de Babel). Excepto para un número «evanescentemente» pequeño de procesos, la inmensa mayoría de estos procesos «hacen muy poca cosa si hacen algo». Así, si encontramos «dos» procesos que hacen algo similar (y que merece ser noticia) casi inevitablemente son un único y el mismo proceso en algún nivel de análisis. Minsky aplica su principio a la aritmética:

Por todo lo dicho, concluyo que cualquier entidad que busca a través de los procesos más simples pronto encontrará fragmentos que no sólo parecen aritméticos sino que *son* aritméticos. No es una cuestión de inventiva o de imaginación, sino tan sólo un hecho en la geografía del universo de la computación, un mundo mucho más sometido a limitaciones que el mundo de las cosas reales (*ibidem*:122).

La cuestión no está evidentemente restringida a la aritmética sino a todas las «verdades necesarias»; lo que los filósofos, desde Platón, han llamado conocimiento *a priori*. Como dice Minsky, «podemos esperar que aparezcan ciertas estructuras *a priori*, casi siempre allí donde un sistema de computación evoluciona por selección desde un universo de posibles procesos» (*ibidem*: 119). A menudo se ha subrayado que la curiosa teoría platónica de la reencarnación y de la reminiscencia, que Platón ofreció como una explicación de la fuente de nuestro conocimiento *a priori*, presenta una semejanza sorprendente con la teoría de Darwin y que esta semejanza es especialmente asombrosa desde nuestra ventajosa posición actual. Es bien conocido que Darwin escribió sobre esta semejanza en uno de sus cuadernos de notas. Comentando la afirmación de que Platón pensó que nuestras «ideas necesarias» surgían de la preexistencia del alma, Darwin anotó: «léase monos donde dice preexistencia» (Desmond y Moore 1991:263).

Por lo tanto, no debe sorprendernos el hallazgo de que los extraterrestres posean la misma inconmovible comprensión que nosotros acerca de que 2+2=4 y otras expresiones aritméticas de este tipo; aunque sí que deberíamos sorprendernos si nos enteramos de que los extra-terrestres utilizan el sistema decimal para expresar sus verdades aritméticas. Nos inclinamos a creer que esta tendencia hacia el sistema decimal es algo así como un accidente histórico, derivado de contar con nuestras manos dotadas de cinco dedos. Pero

supongamos que los extra-terrestres tienen, también, un par de manos, cada una con cinco subunidades. La «solución» de utilizar para contar cualquier cosa que se pueda conseguir es evidentemente obvia, e incluso podría ser incluida en la categoría de los movimientos forzados[46]. No sería demasiado sorprendente averiguar que nuestros alienígenas tienen un *par* de apéndices prehensiles, teniendo en cuenta las buenas razones derivadas de la simetría corporal y la frecuencia de los problemas que se presentan cuando una cosa debe ser manipulada, en relación con otra. El hecho de que haya cinco subunidades en cada apéndice aparece como un fenómeno QWERTY que ha estado profundamente enraizado durante cientos de millones de años; una simple circunstancia fortuita que ha restringido *nuestras* opciones, pero que no debe esperarse que vaya a restringir las suyas. Aunque quizás hayamos subestimado la corrección y la racionalidad de poseer cinco subunidades. Por razones que todavía no hemos desentrañado, esto puede ser, en general, una buena idea y no simplemente algo que nos sorprenda. Por lo tanto, no sería asombroso, después de todo, comprobar que nuestros interlocutores del espacio exterior coinciden en alguna buena idea y cuentan en diez, cientos y miles.

Sin embargo, nos quedaríamos casi sin habla si encontráramos que los extraterrestres utilizan los símbolos que nosotros usamos, los llamados numerales arábigos: «1», «2», «3»… Sabemos que aquí en la Tierra hay magníficas alternativas tales como los numerales arábigos, así como algunas alternativas no viables tales como los numerales romanos, «I», «II», «III», «IV»… Si encontrásemos a los habitantes de otros planetas utilizando numerales arábigos, estaríamos completamente seguros de que no se trata de una coincidencia, sino que debería *haber* una conexión histórica. ¿Por qué? Porque el espacio de las posibles formas numerales en las que no hay *razón* para escoger una entre las otras es astronómicamente grande; la posibilidad de que dos «búsquedas» independientes termine en el mismo lugar es «evanescente».

A menudo, los estudiantes tienen muchas dificultades para comprender la distinción entre números y numerales. Los números son los objetos platónicos abstractos y los numerales son sus nombres. El numeral arábigo «4» y el numeral romano «IV» son sencillamente *nombres* diferentes para la misma cosa, es decir, el *número* 4. (Yo no puedo hablar de un número sin nombrarlo de un modo o de otro, y puedo hablar de Clinton sin usar alguna palabra o palabras que se refieran a él, pero Clinton es un hombre, no una palabra y los números no son símbolos mientras que los numerales sí). Esta es una forma

brillante de demostrar la importancia de la distinción entre numerales y números: precisamente hemos observado que *no* sería sorprendente encontrar que los extraterrestres utilizaran los mismos *números* que nosotros, aunque sería sencillamente increíble que utilizaran los mismos *numerales*.

En un espacio astronómicamente grande de posibilidades, las probabilidades de similitud entre dos elementos escogidos de modo independiente es evanescentemente mínima, *a menos que haya una razón*. Las hay para los números (la aritmética es *verdad* y las variaciones sobre la aritmética no lo son) y no la hay para los numerales (el símbolo &funciona exactamente igual que el símbolo «5» como el nombre del número que sigue al 4).

Supongamos que hemos averiguado que los extraterrestres, como nosotros, utilizan el sistema decimal para sus propósitos más informales, pero recurriendo a la aritmética binaria cuando hacen computación con la ayuda de dispositivos mecánicos protésicos (ordenadores). El uso del 0 y del 1 en sus ordenadores no nos sorprendería (¡suponiendo que los hayan inventado!) dado que existen buenas razones de ingeniería para adoptar el sistema binario, y aunque estas razones no han desaparecido, están probablemente a una llamativa distancia para el tipo medio de pensadores. «No hace falta que seamos científicos espaciales» para apreciar las virtudes del sistema binario.

En general, cabe esperar que los extraterrestres hayan descubierto muchas de las diferentes vías por las que las cosas deben *estar en la vía correcta*. Dondequiera que haya muchos métodos diferentes para despellejar a un gato y ninguno sea bastante mejor que los otros, nuestra sorpresa, si lo hicieran según nuestro método, sería proporcional al número de diferentes métodos que pensamos que existen. Téngase en cuenta que incluso cuando estemos contemplando un gran número de *vías equivalentes*, siempre hay implícito un juicio de valor. Para que podamos reconocer ítems como cosas que cabe incluir en uno de estos enormes grupos, tienen que ser vistos como vías igualmente buenas, como las vías para *realizar la función x*. En este tipo de investigación el pensamiento funcional es sencillamente ineludible; no podemos enumerar las posibilidades sin presuponer un concepto de función. (Ahora podemos comprobar que incluso nuestra formalización deliberadamente aséptica de la Biblioteca de Mendel invocaba presupuestos funcionales; no podemos identificar algo como un *posible genoma* sin pensar en los genomas como cosas que pueden cumplir una particular función dentro de un sistema reproductivo).

De este modo, vuelven a ser principios generales del razonamiento práctico (incluyendo, en una versión más moderna, el *análisis coste-beneficio*) los que pueden llegar a imponerse en todas las formas de vida y en cualquier lugar. Podemos discutir sobre casos particulares, pero no acerca de la aplicabilidad en general de los principios. Las características del diseño, como la simetría bilateral en los seres vivos locomotores, o la boca en el polo anterior, ¿pueden ser explicados como una cuestión de contingencia histórica, o bien como una cuestión de sabiduría práctica? Las únicas cuestiones que pueden ser debatidas o investigadas son las contribuciones relativas y el orden histórico según el cual estas contribuciones han sido realizadas. (Recuérdese que en el fenómeno actual del QWERTY existía una buena *razón* de ingeniería para la elección inicial, aunque las circunstancias que apoyaban dicha razón, hayan desaparecido hace ya mucho tiempo).

El trabajo de diseño —es decir, la llamada elevación en el proceso evolutivo— puede ahora ser definido como el trabajo de descubrir buenas vías para resolver los «problemas que vayan surgiendo». Algunos problemas se presentan al principio, en todos los ambientes, bajo todas las condiciones y en todas las especies. Los iniciales «intentos de solución» de los primeros problemas, puestos en marcha por diferentes especies, crean problemas posteriores. Algunos de estos problemas subsidiarios son creados por otras especies de organismos (que también deben buscarse la vida), y otros problemas subsidiarios son originados por las propias soluciones que las especies dan a sus propios problemas. Por ejemplo, ahora que uno ha decidido —quizá lanzando una moneda al aire— buscar soluciones en *esta* área, se ve sorprendido por el problema B en lugar del problema A, el cual plantea los subproblemas p, q y r, en lugar de los subproblemas x, y y z, y así sucesivamente. ¿Debemos personificar a una especie de esta manera y tratarla como si fuera un agente o un razonador práctico? (Schull 1990 y Dennett 1990b). De modo alternativo, la otra opción sería considerar a las especies como si no fueran agentes, desprovistas de mente, y ubicar la racionalidad en el proceso propio de la selección natural (quizá cómicamente personificado en la llamada madre naturaleza). Recuérdese la agudeza de Crick cuando decía que la evolución es más lista que nosotros. También podemos optar por alejarnos totalmente de estos brillantes modos de expresión, aunque los análisis que hagamos tengan, en cualquier caso, la misma lógica.

Esto es lo que hallamos tras la intuición de que el trabajo de diseño es, de algún modo, un trabajo intelectual. El trabajo de diseño es discernible (en la tipografía ininterpretable de los genomas arrastrados). Sólo si empezamos

imponiéndole *razones*. (En un trabajo anterior he caracterizado a estos genomas como «racionalidades que flotan libremente», una expresión que ha provocado al parecer terror o náusea en muchos lectores, por otra parte bien dispuestos. Confiad en mí; pronto facilitaré algunas formas más aceptables de tratar estas cuestiones).

Así, Paley estaba en lo cierto cuando dijo que explicar el diseño era algo magnífico, pero que también el diseño tomaba posesión de la inteligencia. Todo lo que Paley pasó por alto —y Darwin proporcionó— fue la idea de que esta inteligencia podía ser racionada en *bits* tan mínimos y estúpidos que no podían considerarse como inteligencia y entonces se distribuían, a través del espacio y el tiempo, por una red gigantesca y conectada de procesos algorítmicos. El trabajo debe ser hecho, pero que este trabajo consiga hacerse es, en gran parte, una cuestión de azar, dado que el azar ayuda a determinar qué problemas (y subproblemas y subsubproblemas) llegan a ser «dirigidos» por la maquinaria. Cada vez que encontramos un problema resuelto, podemos preguntar: ¿quién o qué hizo este trabajo? ¿Dónde y cuándo? ¿Ha sido una solución elaborada localmente, o hace mucho tiempo o, de algún modo, tomada prestada (o robada) de alguna otra rama del árbol de la vida? Si exhibe peculiaridades que sólo pueden haber surgido en el curso de la solución de subproblemas en alguna remota rama del árbol de la vida que crece en el espacio de diseño, entonces, exceptuando un milagro o una coincidencia demasiado cósmica para darle crédito, debe ser algún tipo de copia que trasladó este completo trabajo de diseño a su nueva localización.

No existe una única cumbre en el espacio de diseño, ni una única escalinata o escalera con peldaños calibrados, así que no cabe esperar que encontremos una escala que nos permita comparar la cantidad de trabajo de diseño a lo largo de las distantes ramas en desarrollo. Gracias a las vaguedades y digresiones de los diferentes «métodos adoptados», algo que es, en cierto sentido, un problema puede tener soluciones tanto difíciles como fáciles que requieren más o menos trabajo. Hay una famosa historia sobre el matemático y físico John von Neumann (coinventor del ordenador), personaje legendario por su deslumbrante capacidad para realizar cálculos mentales. (Como sucede con las historias más famosas, ésta tiene muchas versiones de las que he escogido la que mejor se adapta a la cuestión que estoy tratando). Un día un colega se le acercó con un rompecabezas para cuya solución existían dos vías, una a través de un laborioso y complicado cálculo, mientras que la otra era una elegante solución de esas que provocan una exclamación del tipo: ¡ajá! Este colega tenía una teoría: en un caso como éste, los

matemáticos lograrían la solución laboriosa, mientras que los físicos (más perezosos pero más inteligentes) harían una pausa y encontrarían la solución rápida y fácil. ¿Qué solución encontraría Von Neumann? Este era el rompecabezas: dos trenes, separados por roo millas, se aproximan uno al otro, marchando uno a 30 millas por hora y el otro a 20 millas por hora. Un pájaro que vuela a no millas por hora comienza desde el tren A (cuando están ambos trenes separados por 100 millas), vuela hacia el tren B, da la vuelta y vuela hacia el tren A que se aproxima, y así sucesivamente hasta que los trenes colisionan. ¿A qué distancia volaba el pájaro cuando ocurre el choque entre ambos trenes? «Doscientas cuarenta millas», respondió Von Neumann, casi instantáneamente. «¡Maldición!», replicó su colega. «Yo predije que escogerías la solución laboriosa». «¡Ay!», exclamó Von Neumann avergonzado golpeando su frente. «¡Hay una solución fácil!». (Insinuación: ¿cuánto tiempo transcurre hasta que los trenes colisionan?).

Los ojos son el ejemplo estándar de un problema que ha sido resuelto muchas veces, aunque los ojos que pueden mirar justamente lo mismo (y ver lo mismo) cabe que hayan sido conseguidos mediante proyectos de I+D que incluyan diferentes cantidades de trabajo, gracias a las peculiaridades históricas de las dificultades con las que tropezaron a lo largo del camino. Y las criaturas que no tienen ojos no son ni mejores ni peores en ninguna escala absoluta del diseño: su linaje nunca ha tenido necesidad de resolver este problema. Es esta misma variabilidad en la *fortuna* de los diferentes linajes la que hace imposible definir un simple punto de apoyo como pedía Arquímedes, desde el cual el proceso en su globalidad pueda ser medido. ¿Es progreso que tengamos que trabajar en una ocupación laboral extraordinaria para pagar el elevado precio de un mecánico que hemos contratado para reparar nuestro coche cuando se estropea, porque es demasiado complejo para nosotros repararlo del modo que utilizamos para arreglar nuestro viejo cacharro? ¿Quién lo dice? Algunos linajes quedan atrapados (o son lo bastante afortunados como para descarriarse, puede elegirse la expresión que se prefiera) en un camino en el espacio de diseño en el que la complejidad engendra complejidad, en una «carrera de armamento» de diseño competitivo. Otros son los bastantes afortunados (o bastante desafortunados, elíjase lo que se prefiera) para alcanzar desde el principio una solución relativamente simple para el problema de la vida y, habiéndola descubierto hace mil millones de años, desde entonces no han tenido nada que hacer, en lo que se refiere a su diseño. Nosotros, seres humanos, y como tales, criaturas complicadas, tendemos a valorar la complejidad, aunque ésta puede ser muy

bien una preferencia estética ligada a nuestro linaje; otros linajes pueden ser tan felices como lo son las almejas con su cuota de simplicidad.

3. *La unidad del espacio de diseño*

> La formación de los diferentes lenguajes y de las distintas especies y las pruebas de que ambos se han desarrollado a través de un proceso gradual, son curiosamente una misma cosa.
>
> Charles Darwin, *El origen del hombre*

No habrá pasado inadvertido que mis ejemplos en este capítulo han oscilado entre el dominio de los organismos o diseño biológico, por un lado, y el dominio de los artefactos humanos —libros, problemas resueltos y triunfos de la ingeniería— por otro. Es evidente que esto ha sido así por diseño, no por accidente. Con la ayuda de estos ejemplos hemos preparado el escenario y hecho acopio de la suficiente munición para disparar una gran salva: *sólo hay un espacio de diseño y todas aquellas cosas que son reales en este espacio están unidas a todas las demás*. Es evidente que no necesito añadir que fue Darwin quien nos transmitió esta idea, se diera o no cuenta, en aquel momento, de lo que estaba proponiendo.

Ahora pretendo volver atrás sobre el territorio del que ya nos hemos ocupado para resaltar la evidencia de esta declaración y extraer algunas implicaciones adicionales de esta idea, y los fundamentos que permiten darle crédito. Creo que las similitudes y las continuidades son muy importantes, aunque en los próximos capítulos también señalaremos algunas diferencias importantes entre las partes del mundo diseñado hechas por el hombre y aquellas partes que han sido creadas sin el beneficio de ese tipo de inteligencia que los artífices humanos aplicamos en la resolución de un problema, que es una inteligencia capaz de concentrarse localmente y, al mismo tiempo, previsora.

Al principio hemos subrayado que la Biblioteca de Mendel (bajo la forma de volúmenes impresos con las letras A, C, G y T) está contenida en la Biblioteca de Babel, pero también debemos recordar que una gran parte, al menos, de la Biblioteca de Babel (¿qué parte?; véase el capítulo 15) está, a su vez, contenida en la Biblioteca de Mendel, debido a que *nosotros* estamos en la Biblioteca de Mendel (están nuestros genomas y están los genomas de todas las formas de vida de las que nuestras vidas dependen). La Biblioteca de Babel describe un aspecto de nuestro *fenotipo extendido* (Dawkins 1982). Es

decir, del mismo modo que las arañas tejen redes y los castores construyen presas, nosotros hacemos (entre otras muchas cosas) libros. No podemos valorar la viabilidad del genoma de la araña sin tener en cuenta la red que es parte de su equipamiento normal, y no podemos evaluar la viabilidad de nuestros genomas (no podremos durante mucho más tiempo) sin reconocer que somos una especie con cultura, de la que una parte representativa se encuentra en forma de libros. No hemos sido precisamente diseñados, somos diseñadores, y todos nuestros talentos como diseñadores, así como nuestros productos, deben emerger, *de una forma o de otra*, sin milagros, de los ciegos y mecánicos procesos de los mecanismos de Darwin. ¿Cuántas grúas, colocadas unas encima de otras, han sido utilizadas para conseguir desde las primeras indagaciones de diseño en los linajes procariotas hasta las investigaciones matemáticas de los profesores de Oxford? Esta es la cuestión que pone sobre la mesa el pensamiento darwiniano. La resistencia viene de aquellos que piensan que entre los procariotas y los más hermosos tesoros de nuestras bibliotecas debe haber algunas discontinuidades en algún lugar, algunos «ganchos celestes» o algunos momentos de creación especial, o algún otro tipo de milagro.

Esta puede ser una cuestión que cabe examinar desde muy diversos puntos de vista en el resto del libro. Pero ya hemos encontrado una variedad de profundos paralelismos, ejemplos en los que los mismísimos principios, las mismísimas estrategias de análisis de la inferencia se aplican en ambos dominios. Hay muchos más puntos de partida de donde pueden proceder.

Consideremos, por ejemplo, la utilización pionera por Darwin de un cierto tipo de inferencia histórica. Como ha subrayado Stephen Jay Could (por ejemplo 1977a y 1980a), son las imperfecciones, los curiosos fallos que se quedan a muy corta distancia de lo que sería la perfección de un diseño, la mejor evidencia de un proceso histórico de «descenso con modificaciones»; son también la mejor evidencia del proceso de la copia, en lugar de una reinvención independiente del diseño en cuestión. Ahora podemos comprender mejor por qué ésta es una evidencia tan buena. Las probabilidades contra la existencia de dos procesos independientes que lleguen a coincidir en la misma región del espacio de diseño son enormemente grandes, a menos que el elemento de diseño en cuestión, obviamente correcto, sea un movimiento forzado en ese espacio. La perfección será independientemente alcanzada una y otra vez, sobre todo si es obvia. Las versiones idiosincráticas de la casi perfección son un premio inútil del proceso de la copia. En la teoría evolucionista tales caracteres se llaman

homologías: caracteres que son similares, no por razones funcionales sino porque son copias. El biólogo Mark Ridley ha señalado que «muchos de los argumentos que, a menudo, son presentados como independientes para la teoría de la evolución se pueden reducir al esquema general del argumento de la homología» y reduce el argumento a su esencia de la siguiente forma:

Los huesos del oído de los mamíferos son un ejemplo de homologías. Estos huesecillos son homólogos de algunos de los huesos mandibulares de los reptiles. Los huesos del oído de los mamíferos no han sido formados a partir de los mismos huesos que constituyen la mandíbula de los reptiles; pero, de hecho, lo son… El hecho de que las especies compartan homologías es un argumento a favor de la evolución, porque si hubieran sido creadas separadamente, *no habría razón para que* mostraran similitudes homologas (Mark Ridley 1985:9). (La cursiva es mía).

Así sucede también en la biosfera, y dentro de la esfera de la cultura con el plagio, con el espionaje industrial y con el honrado trabajo de la *recensión de textos*.

He aquí una curiosa coincidencia: mientras Darwin luchaba por hacer comprender con claridad el característico método darwiniano de la inferencia, algunos de sus colegas Victorianos, en Inglaterra y especialmente en Alemania, ya habían perfeccionado la misma estrategia valiente e ingeniosa de la inferencia histórica en el dominio de la *paleografía* o de la *filología*. He aludido varias veces en este libro a las obras de Platón, pero es un «milagro» que la obra de Platón haya sobrevivido para que nosotros podamos leerla en cualquier versión. Todos los textos de sus *Diálogos* estuvieron perdidos durante cerca de mil años. Cuando volvieron a aparecer, en los albores del Renacimiento, bajo la forma de varias copias de copias de copias, deterioradas, dudosas y parciales, realizadas por quienes sabían dónde se encontraban, se puso en marcha durante quinientos años una cuidadosa investigación de las copias, que trataba de «limpiar el texto» y establecer unas conexiones informativas apropiadas con las fuentes originales, las cuales, naturalmente, habrían estado en las propias manos de Platón o en la mano del escriba a quien se le había dictado. Es de suponer que los originales se convirtieron en polvo hace ya mucho tiempo. (Hoy se dispone de fragmentos de papiros con textos platónicos y estos *bits* de texto pueden ser aproximadamente contemporáneos del propio Platón, aunque los papiros no hayan desempeñado un papel importante en la investigación realizada, ya que han sido recuperados muy recientemente).

La tarea con la que se enfrentaron los eruditos era intimidadora. Obviamente se encontraron muchas «corrupciones» en las diferentes copias supervivientes (denominadas «testigos») y estas corrupciones o errores tenían

que ser solucionadas, pero también había muchos pasajes confusos —o emocionantes— de dudosa autenticidad, sin que fuera posible preguntarle al autor. ¿Cómo podían distinguirse apropiadamente los errores? Las corrupciones, más o menos, podían ser ordenadas respecto a su obviedad de la siguiente manera: 1) errores tipográficos, 2) errores gramaticales, 3) expresiones estúpidas o confusas, 4) fragmentos (*bits*) que, desde el punto de vista estilístico o doctrinal, no concordaban con el resto del texto platónico. En los tiempos de Darwin, los filólogos que dedicaban su vida profesional a recrear la genealogía de sus «testigos» no sólo habían desarrollado unos elaborados —para su tiempo— y rigurosos métodos de comparación, sino que habían tenido éxito en la extrapolación de la totalidad de los linajes de copias de copias, y pudieron deducir muchos hechos curiosos acerca de las circunstancias históricas de su nacimiento, de su reproducción y, eventualmente, de su muerte. Mediante un análisis de los modelos de los errores compartidos y no compartidos en los documentos existentes (los tesoros de pergamino cuidadosamente conservados en la Biblioteca Bodleian de Oxford, en París, en la Biblioteca Nacional de Viena, en el Vaticano y en cualquier otro lugar), los investigadores fueron capaces de deducir hipótesis acerca de cuántas copias diferentes habían sido hechas aproximadamente, cuándo y dónde algunas de ellas habían sido realizadas y qué copias «testigos» han compartido recientes ancestros y cuáles no.

A veces la arriesgada deducción que nace de sus trabajos es similar a la de Darwin: un grupo concreto de errores en los manuscritos, no corregidos y vueltos a copiar en todos los manuscritos descendientes de un particular linaje, se debía casi siempre al hecho de que el escriba (copista) que escribía al dictado no pronunciaba el griego de la misma forma que lo hacía el lector del texto y, en consecuencia, ¡entendía incorrectamente, en muchas ocasiones, un determinado fonema! Estas claves, junto con la evidencia procedente de otras fuentes de la historia de la lengua griega, pudieron incluso haber sugerido a los eruditos en qué monasterio, en una isla griega o en la cima de una montaña, y en qué siglo debía haber tenido lugar la escena en la que se produjo este conjunto de mutaciones, incluso a pesar de que el documento escrito entonces sobre pergamino haya sucumbido hace ya mucho tiempo a la segunda ley de la termodinámica y se hubiera convertido en polvo[47].

¿Aprendió algo Darwin de los filólogos? ¿Ha reconocido algún filólogo que Darwin había reinventado una de sus ruedas? El mismo Nietzsche fue uno de esos magníficos estudiosos de textos antiguos, y uno de los muchos pensadores alemanes que habían estado implicados en el *boom* de Darwin,

pero, hasta lo que yo conozco, nunca se dio cuenta de que existiera un parentesco entre el método de Darwin y los de sus colegas. En los últimos años, a Darwin le llamó la atención la curiosa similitud entre sus argumentos y los de los filólogos que estudian la genealogía de los lenguajes (no, como en el caso de los estudiosos de Platón, la genealogía de textos específicos). En *El origen del hombre*, Darwin señalaba explícitamente el uso compartido de la distinción entre homologías y analogías, que podía ser debido a una evolución convergente: «Hemos encontrado en lenguajes distintos llamativas homologías debidas a una comunidad en el descenso y analogías debidas a un proceso similar de formación».

Las imperfecciones o errores son justamente casos especiales en la variedad de señales que nos hablan en voz alta —e intuitivamente— de una historia compartida. El papel del azar en los retorcidos caminos seguidos en un fragmento de un trabajo de diseño puede producir el mismo efecto de señal sin crear un error. Un caso ejemplar: en el año 1988 le fue enviada a Otro Neugebauer, el gran historiador de la astronomía, una fotografía de un fragmento de un papiro griego con unos cuantos números en una columna. Quien se la enviaba, un estudioso de la antigüedad clásica, no conseguía descubrir la clave que le permitiese conocer el significado de ese fragmento de papiro, y le preguntaba a Neugebauer si se le ocurría alguna idea para lograrlo. A sus ochenta y nueve años, Neugebauer comprobó, línea por línea, las diferencias entre los números, hallando sus límites máximos y mínimos, y determinó que este papiro ¡era una traducción de parte de la «columna G» de una tabla babilónica con escritura cuneiforme en la que había escrito un «sistema A» de Babilonia para las efemérides lunares! (Una efeméride es, como el *Almanaque Náutico*, un sistema tabular para computar la localización de un cuerpo celeste en cualquier momento en un período concreto). ¿Cómo *pudo* Neugebauer llegar a esta deducción al estilo de Sherlock Holmes? Elemental: lo que estaba escrito en griego (una secuencia de números sexagesimales, no decimales) fue reconocido por él como parte —¡columna G!— de un cálculo extraordinariamente preciso de la localización de la Luna que había sido elaborado por los babilonios. Hay muchos métodos diferentes para calcular una efeméride y Neugebauer sabía que alguien que trabaja sus propias efemérides de modo independiente, usando su propio sistema, no llegaría exactamente a los mismos números, aunque hubiesen estado tan cercanos. El sistema B de los babilonios era excelente, así que el diseño había sido afortunadamente conservado, traducido, con todas sus mínimas particularidades (Neugebauer 1989)[48].

Neugebauer fue un gran investigador, pero nosotros podemos probablemente llevar a cabo una proeza deductiva semejante si seguimos sus pasos. Supongamos que se nos envía una fotocopia del texto de la figura y se nos hace la misma pregunta. ¿Qué significa? ¿De dónde procede este texto?

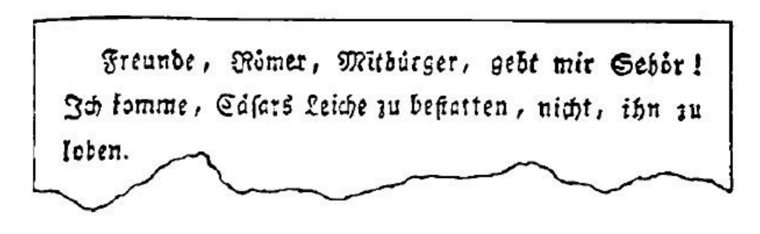

Freunde, Römer, Mitbürger, gebt mir Gehör! Ich komme, Cäsars Leiche zu bestatten, nicht, ihn zu loben.

Figura 6.1

Antes de leerlo, probemos. ¡Podemos probablemente hacer una deducción, incluso si no sabemos realmente cómo leer el antiguo tipo de letra alemán conocido como *Fraktur*, e incluso sin saber alemán! Miremos de nuevo, más de cerca. ¿Hemos averiguado algo? ¡Qué habilidad tan impresionante! Neugebauer pudo contar con su babilónica columna G, pero nosotros rápidamente determinamos, ¿no es así?, que este fragmento debe ser parte de una traducción al alemán de algunas líneas de una tragedia isabelina (*Julio César*, III, 11, 79-80, para ser exactos). ¡Una vez que reflexionamos sobre esto nos damos cuenta de que difícilmente podía ser otra cosa! Las probabilidades contra *esta* particular secuencia de letras alemanas que se encuentran alineadas juntas son astronómicamente enormes bajo cualquier otra circunstancia. ¿Por qué? ¿Cuál es la particularidad que marca a esta serie de símbolos?

Nicholas Humphrey planteó esta pregunta de manera muy gráfica mediante una versión más drástica: si nos obligaran a «consignar al olvido» una de las siguientes obras maestras, ¿cuál elegiríamos?: ¿los *Principia* de Newton, los *Cuentos de Canterbury* de Chaucer, el *Don Giovanni* de Mozart o la Torre Eiffel? «Si la elección fuera obligada», respondió Humphrey,

> tengo pocas dudas de lo que yo haría: serían los *Principia* de Newton los que irían al olvido. ¿Por qué? Porque de todas aquellas obras, la de Newton es la única que es reemplazable. Muy simple; si Newton no la hubiera escrito, otro lo hubiera hecho, probablemente en pocos años… Los *Principia* son realmente un monumento glorioso de la inteligencia humana, la Torre Eiffel una hazaña relativamente menor de la ingeniería romántica; el hecho es que mientras la Torre Eiffel siguió *su* camino, Newton siguió el camino de Dios (Humphrey 1987).

Newton y Leibniz entablaron una famosa discusión acerca de cuál de los dos fue el inventor del cálculo infinitesimal; podemos imaginarnos fácilmente a Newton discutiendo con un contemporáneo sobre cuál de los dos tenía la prioridad del descubrimiento de las leyes de la gravitación. Pero si Shakespeare no hubiera vivido nunca, nadie hubiera escrito sus obras de teatro y sus poemas. «C. P. Snow, en su libro *Las dos culturas* (1963), elogia los grandes descubrimientos de la ciencia bajo el apelativo de "el científico Shakespeare". Pero, de algún modo. Snow esta ha completamente equivocado. Las obras de teatro de Shakespeare eran las obras de teatro de Shakespeare y no de otro autor. Por el contrario, los descubrimientos científicos pertenecen, en último término, a uno en particular» (Humphrey 1987). Intuitivamente la diferencia estriba en la diferencia entre el descubrimiento y la creación, aunque ahora tenemos una manera mejor de comprobar esta diferencia. Por un lado, está el trabajo de diseño que está orientado hacia el mejor movimiento o el movimiento forzado, el cual puede ser visto (retrospectivamente, al menos) como la única localización favorecida en el espacio de diseño accesible desde muchos puntos de inicio y a través de muy diferentes caminos; por otro lado, está el trabajo de diseño de excelencia que depende mucho más de la explotación o amplificación de las muchas contingencias de la historia que moldean su trayectoria, una trayectoria para la que el eslogan de la compañía de autobuses es una declaración excesivamente modesta: llegar allí es bastante más que la mitad de la diversión.

En el capítulo 2 hemos visto que incluso el algoritmo de divisiones largas puede beneficiarse de la aleatoriedad o de una idiosincrasia arbitraria: escoger un dígito al azar (o nuestro dígito favorito) y comprobar si es el «correcto». Pero las opciones idiosincráticas reales escogidas en tanto se prosigue la búsqueda del equilibrio de la operación no dejan huella en la respuesta final, la respuesta correcta. Otros algoritmos pueden incorporar las opciones aleatorias en la estructura de sus productos finales. Piénsese en un algoritmo para escribir poemas —o, si se desea, en un algoritmo para escribir coplillas burlescas— que comienza: «escoja un nombre al azar en el diccionario…». Este proceso de diseño puede producir algo que está definitivamente bajo control de calidad —presión de selección— pero que, a pesar de todo, presenta signos inequívocos de la particular historia de su creación.

El contraste establecido por Humphrey es agudo, pero su brillante modo de esquematizarlo puede ser engañoso. La ciencia, y no las artes, está implicada en viajes —a veces en carreras— con destinos definidos: alcanzar

soluciones para problemas específicos en el espacio de diseño. Pero los científicos son muy cautelosos, como bastantes artistas, con las rutas que han de tomar, y en consecuencia quedan consternados ante la idea de tener que abandonar el trabajo real de Newton y mantener su destino (al que posiblemente alguien nos hubiera conducido, en cualquier caso). Ellos se preocupan de las trayectorias reales debido a que los métodos utilizados pueden ser reutilizados para otros viajes; los buenos métodos son grúas, que pueden ser tomadas como préstamo, con reconocimiento, y emplearse para hacer elevaciones en otras partes del espacio de diseño. En el caso extremo, la grúa desarrollada por un científico puede ser de bastante más valor que el hecho de la concreta elevación cumplida por aquélla, es decir, alcanzar el destino. Por ejemplo, una prueba que produce un resultado trivial puede, no obstante, ser pionera de un nuevo método matemático «grúa» de gran valor. Los matemáticos conceden un gran valor al hecho de probar con una prueba más simple y elegante algo que ellos ya han comprobado, es decir, a lograrlo con una grúa más eficiente.

En este contexto, la filosofía puede considerarse ubicada a medio camino entre la ciencia y las artes. Es bien conocido que Ludwig Wittgenstein consideraba que en la filosofía el proceso —la argumentación y el análisis— era más importante que el producto, las conclusiones alcanzadas y la teorías defendidas. Aunque esta opinión es muy discutida —y estoy de acuerdo— por muchos filósofos que aspiran a resolver problemas reales y no caer en una suerte de logoterapia interminable, se debe admitir, por ejemplo, que no deseamos consignar nunca al olvido la famosa experiencia del pensamiento de Descartes expresada como «cogito ergo sum», a pesar de que nadie aceptaría hoy sus conclusiones: es precisamente una «bomba» de intuición demasiado elegante, incluso si todas sus intuiciones fueran falsas (Dennett 1984:18).

¿Por qué no podemos solicitar los derechos de autor por una multiplicación de dos números correctamente realizada? Porque cualquiera puede hacerlo. Es un movimiento forzado. La misma razón puede aplicarse a cualquier hecho simple que no necesita de un genio para ser descubierto. ¿Por qué los creadores de tablas de otras rutinas muy trabajosas, repletas de datos, las protegen de copistas nada escrupulosos? A veces preparan trampas. Me han contado, por ejemplo, que los editores de *Who's Who* han tratado el problema de los competidores simplemente escondiendo todos los datos que obtienen con más dificultad y publicando sus propias enciclopedias biográficas, en las que insertan unas cuantas entradas falsas. ¡Podemos estar

seguros de que si alguno de aquellos datos aparece en las páginas de un competidor, no se trata de una coincidencia!

En la más amplia perspectiva de la totalidad del espacio de diseño, el delito del plagio puede ser definido como el *robo de una grúa*. Alguien o algo ha hecho un trabajo de diseño por medio del cual ha creado algo que es útil para trabajos de diseño posteriores y, en consecuencia, puede tener valor para alguien o algo embarcado en un proyecto de diseño. En el mundo de la cultura, donde la transmisión de los diseños de agente a agente es facilitada por los medios de comunicación, la adquisición de diseños desarrollados en otros «talleres» es un hecho frecuente, casi el marco que define la evolución cultural (que sera el tema del capítulo 12). Los biólogos han asumido habitualmente que dichas transacciones eran imposibles en el mundo de la genética (hasta la aparición de la ingeniería genética). De hecho, puede decirse que este ha sido el dogma oficial. Descubrimientos recientes sugieren lo contrario, aunque sólo el tiempo lo dirá; ningún dogma es derribado y muerto sin lucha. Por ejemplo, Marilyn Houck (Houck y otros 1991) ha hallado evidencias de que, hace cuarenta años, en Florida o en América Central, un pequeño ácaro que se alimenta de las moscas de la fruta perforó el huevo de una mosca de la especie *Drosophila willistoni* y, en el proceso, recogió algo del ADN característico de esta especie y lo transmitió inadvertidamente al huevo de una mosca (salvaje) de la especie *Drosophila melanogaster*. Este accidente podría explicar la súbita explosión en la *Drosophila melanogaster* de un particular elemento del ADN común en la de las moscas de la fruta, aunque previamente desconocido en las poblaciones de la *D. melanogaster*. Marilyn Houck pudo añadir: ¿qué otra cosa puede explicar este hallazgo? Seguramente se trata de un plagio entre especies.

Otros investigadores están buscando otros posibles vehículos para un veloz viaje de diseño en el mundo de la genética natural (como opuesto a la genética artificial). Si los encuentran, serán fascinantes, aunque sin duda raras excepciones del modelo ortodoxo: transmisión genética del diseño sólo a través de las cadenas de descendientes directos[49]. Estamos inclinados, como hemos llamado la atención, a contrastar esta característica de manera tajante con lo que encontramos en el espontáneo mundo de la evolución cultural, pero incluso en este mundo podemos detectar una poderosa dependencia de la combinación entre azar y copia.

Pensemos en todos los libros admirables de la Biblioteca de Babel que nunca serán escritos, aunque el proceso que crearía a cada uno de ellos no implicaría violación o limitación de las leyes de la naturaleza. Consideremos

algún libro de la Biblioteca de Babel que a cualquiera de nosotros le gustaría escribir —y que sólo él podría escribir—, por ejemplo, la historia autobiográfica de su infancia, expresada con un lenguaje poético, que provocaría lágrimas y risas en todos sus lectores. Sabemos que hay un número astronómicamente grande de libros con estas características en la Biblioteca de Babel y cada uno compuesto con sólo un millón de pulsaciones de teclas. Al perezoso ritmo de quinientas pulsaciones diarias, el proyecto global no debería tardar más de seis años, incluyendo unas generosas vacaciones. Pues bien, ¿qué nos detiene? Tenemos dedos para trabajar y todas las teclas en el ordenador pueden ser pulsadas de modo independiente.

No nos detiene nada. Es decir, no es necesario que haya fuerzas identificables o leyes de la física o la biología o la psicología, o asombrosas discapacidades provocadas por circunstancias identificables (tales como un hacha clavada en el cerebro o un agresor que nos apunte con una pistola). Hay una cantidad de libros astronómicamente grande que no llegaremos a escribir nunca «por ninguna razón» especial. Gracias a los miles y miles de idas y venidas de nuestra vida hasta la fecha, no hemos encontrado una ocasión propicia para componer aquellas secuencias de pulsaciones de teclas.

Si queremos obtener alguna perspectiva —seguramente limitada— sobre los modelos que se han tenido en cuenta para crear nuestra propia disposición como autor, tenemos que considerar la transmisión hasta nosotros del diseño de los libros que hemos leído. Los libros que realmente han llegado a existir en las bibliotecas del mundo dependen por completo no precisamente de la herencia biológica de sus autores, sino de los libros que han aparecido antes que ellos. Esta dependencia ha sido condicionada por coincidencias o accidentes en cada recodo del camino seguido. Revisemos mi bibliografía para descubrir las principales líneas de la genealogía de este libro. He estado leyendo y escribiendo sobre la evolución desde que era un estudiante, pero si Doug Hofstadter en 1980 no me hubiera estimulado a leer el libro de Dawkins *El gen egoísta*, probablemente no hubiera comenzado a conjugar algunos intereses y hábitos de lectura que han sido los que más han contribuido a dar forma a este libro. Y si *The New York Review of Books* no le hubiese solicitado a Hofstadter la revisión de mi libro *Brainstorms* (1978) no se le hubiera ocurrido la brillante idea de que colaborara con él en el libro *The Mind's 1* (1981) y entonces no hubiéramos tenido la oportunidad de recomendarnos libros, circunstancias que me llevaron a Dawkins y de ahí en adelante. Aunque si yo hubiera leído los mismos libros y artículos siguiendo otros caminos, en un orden diferente, no hubiera estado condicionado a seguir

exactamente la misma vía para realizar esta lectura y, en consecuencia, habría sido muy improbable que escribiera —y revisara una y otra vez— la retahíla de símbolos que ahora tiene ante sí el lector.

¿Podemos medir esta transmisión de diseño como cultura? ¿Existen unidades de transmisión cultural análogas a los genes de la evolución biológica? Dawkins [1976] propone su existencia y les da un nombre: *memes*. Como los genes, los memes se supone que son replicadores en un medio diferente, pero sujetos a los mismos principios de la evolución de los genes. La idea de que pueda existir una teoría científica, la memética, estrechamente paralela a la genética, les parece a muchos un absurdo pero, al menos, una gran parte de este escepticismo está basado en una mala interpretación. Esta es una idea controvertida, a la que prestaremos atención en el capítulo 12; mientras tanto, la dejaremos a un lado y usaremos el término como una palabra que tendremos a mano para aplicarla a un hecho cultural relevante (*memorable*), algo con bastante diseño como para merecer ser conservado, robado o replicado.

La Biblioteca de Mendel (o su gemela Biblioteca de Babel, después de todo cada una de ellas está contenida en la otra) es tan buena como modelo aproximado del espacio universal del diseño que es necesario reflexionar sobre ella. En los últimos cuatro mil millones de años, poco más o menos, el árbol de la vida ha crecido zigzagueando a través de este enorme espacio multidimensional, ramificándose y floreciendo con fecundidad virtualmente inimaginable, pero, sin embargo, tratando de llenar sólo una pequeña porción evanescente, de este espacio de lo posible con diseños reales[50]. De acuerdo con la peligrosa idea de Darwin, todas las *posibles* exploraciones del espacio de diseño están conectadas. No sólo nuestros hijos y los hijos de nuestros hijos, sino todos los cerebros de nuestros hijos y los cerebros de los hijos de éstos deben crecer a partir del depósito común de elementos de diseño, genes y memes, que desde hace tanto tiempo han sido acumulados y conservados por la inexorable elevación de los algoritmos, las rampas y las grúas, y las grúas sobre las grúas, de la selección natural y sus productos.

Si esto es correcto, entonces todos los logros de la cultura humana —lenguaje, arte, religión, ética, la ciencia misma— son en sí mismos artefactos (de artefactos de artefactos…) del mismo proceso fundamental que ha desarrollado la bacteria, los mamíferos y el *Homo sapiens*. No ha habido una creación especial del lenguaje y ni el arte ni la religión tienen literalmente

inspiración divina. Si no existen «ganchos celestes» que sean necesarios para hacer una alondra, tampoco hay «ganchos celestes» para componer una oda a un ruiseñor. Ningún *meme* es una isla.

La vida y todas sus glorias están así unidas bajo una perspectiva singular, aunque a algunos esta visión les parezca detestable, yerma y odiosa. Desean clamar contra ella y, por encima de todo, desean ser magníficas excepciones. Si no lo es el resto, al menos ellos están creados por Dios a su imagen y semejanza, o, si no son religiosos, desean «ganchos celestes» para ellos mismos. Desean, de algún modo, ser *fuentes intrínsecas* de inteligencia o Diseño, no «meros» artefactos del mismo proceso insensato que ha producido el resto de la biosfera.

Muchas cosas están en juego. Antes de que volvamos, en la tercera parte, a examinar con detalle las implicaciones de la extensión ascendente del ácido universal a través de la cultura humana, necesitamos asegurar el campo base, considerando una serie de profundos retos al pensamiento darwiniano dentro de la propia biología. En el proceso, nuestra visión de lo intrincado y poderoso de las ideas subyacentes se potenciará.

PARTE II
El pensamiento de Darwin en biología

La evolución es un cambio desde una total similitud de la que no se puede hablar en absoluto mediante continuas acciones de conjunción y acciones de algo más[2t].

William James, *Lecture Notes 1880-1897*

Nada en biología tiene sentido si no es a la luz de la evolución.

Thedosius Dobzhansky,
«Nothing in Biology' Makes Sense Except in the Light of Evolution»

CAPÍTULO 7
Cebando la bomba de Darwin

I. *Retrocediendo más allá de la frontera de Darwin*

> Y Dios dijo: «Que brote en la tierra hierba verde, plantas que den semilla, árboles frutales que produzcan frutos según su especie y cuya semilla esté en ellos sobre la tierra». Y así fue.
>
> Brotó, pues, en la tierra hierba verde, plantas que tenían semillas según su especie, y árboles que producían frutos y cuya semilla se hallaba en ellos según su especie. Y vio Dios que era bueno.
>
> Génesis

¿A partir de qué tipo de semilla pudo iniciar su crecimiento el árbol de la vida? En la actualidad se ha establecido, más allá de toda duda razonable, que toda la vida sobre la Tierra fue producida por un proceso de generación que se va ramificando. En gran parte gracias a Darwin, este proceso es un ejemplo de hecho científico tan seguro como lo es la redondez de la Tierra. Pero ¿cómo comenzó el proceso global en su primera localización? Como hemos visto en el capítulo 3, Darwin no sólo empezó por la mitad, sino que, con cautela, se abstuvo de extender su ya público propio pensamiento hacia atrás, es decir, al inicio, o sea, a los orígenes de la vida y a sus condiciones previas. Cuando los que mantenían correspondencia con Darwin lo presionaban, en privado tenía algo más que añadir. En una famosa carta conjeturaba que era completamente posible que la vida hubiera comenzado en una «pequeña charca caliente», aunque no tenía detalles que ofrecer acerca de la probable receta para esta primigenia sopa preorgánica. Y en una respuesta a Asa Gray, como ya hemos visto (p. 99), dejó abierta de par en par la puerta a la posibilidad de que las leyes que gobiernan estos movimientos en la agitada Tierra estuvieran diseñadas presumiblemente por Dios.

Su reticencia a este respecto era sabia desde varios puntos de vista. En primer lugar, nadie conocía mejor que él la importancia de anclar una teoría revolucionaria en el suelo rocoso de los hechos empíricos, sabiendo que sólo podía especular, con una débil esperanza, en su tiempo, de alcanzar algún

sustantivo *feedback*. Después de todo, como ya hemos visto, Darwin no poseía el concepto mendeliano del gen ni de la maquinaria molecular que lo sustentaba. Darwin estaba dotado de una capacidad deductiva atrevida, aunque también sabía cuándo no se encontraba en posesión de las suficientes premisas para seguir adelante. A esto se añadía su preocupación por su amada esposa, Emma, que deseaba asirse desesperadamente a sus creencias religiosas al ver la amenaza que se vislumbraba en los trabajos de su marido. Sin embargo, su resistencia a penetrar aún más en este peligroso territorio, al menos en público, fue superior a sus consideraciones por los sentimientos de su esposa. Estaba en juego una consideración ética más amplia, que ciertamente gozaba del aprecio de Darwin.

Se ha escrito bastante acerca de los dilemas morales a los que los científicos se enfrentan cuando el descubrimiento de un hecho potencialmente peligroso coloca su amor por la verdad en desacuerdo con su preocupación por el bienestar de los demás. ¿Bajo qué condiciones, si hay alguna, estarían los científicos obligados a ocultar la verdad? Estos pueden ser verdaderos dilemas, con consideraciones poderosas y difíciles de probar por ambas partes. Pero no existe, en absoluto, controversia acerca de lo que deben ser las obligaciones morales de un científico (o de un filósofo) en lo que se refiere a sus especulaciones. La ciencia, a menudo, no avanza por la acumulación metódica de hechos demostrables; el «filo cortante» del avance casi siempre está compuesto por varios filos rivales, que compiten afiladamente y que son descaradamente especulativos. Muchas de estas especulaciones pronto se comprueba que son espurias, por muy atractivas que hayan sido al comienzo, y estos necesarios productos de desecho de la investigación científica deben ser considerados como potencialmente peligrosos, como todos los desechos de laboratorio. No obstante, conviene tener en cuenta su impacto ambiental. Si su incorrecta comprensión por el público puede llegar a producir sufrimiento —por gente peligrosa que lleva a cabo acciones peligrosas, o bien socava su fidelidad a algún principio socialmente deseable o a un credo—, los científicos deben ser especialmente prudentes en el modo de proceder, escrupulosos respecto a calificar las especulaciones como lo que son, y deben limitar la retórica de la persuasión a sus propios objetivos.

Pero las ideas, al contrario que los humos tóxicos y los residuos químicos, son casi imposibles de someter a cuarentena, sobre todo cuando se refieren a temas de continuada curiosidad humana; así que mientras no hay controversia, en absoluto, acerca del principio de responsabilidad, ha habido

escaso acuerdo, entonces y ahora, respecto a cómo honrarlo. Darwin hizo lo mejor que podía haber hecho: guardar para él mismo sus especulaciones.

Nosotros lo podemos hacer mejor. En nuestro tiempo, la física y la química de la vida se comprenden en sus más brillantes detalles, de modo que pueden ser deducidas bastante más cosas sobre las necesarias y (quizá) suficientes condiciones para la vida. Las respuestas a las grandes cuestiones pueden conducir todavía a la especulación, aunque podemos indicar que se trata de especulaciones y advertir que éstas pueden ser confirmadas o refutadas. Ya no es cuestión de continuar con una política de reticencia respecto a Darwin; demasiados pasteles muy atractivos se encuentran ya al descubierto. Todavía no sabemos con exactitud *cómo* tomar en serio todas estas ideas, pero gracias a la asegurada cabeza de puente que Darwin ha establecido en la biología, conocemos lo que podemos y lo que debemos.

Sorprende un poco que Darwin no acertara con un mecanismo apropiado de la herencia. ¿Es que debemos suponer que su actitud debería haber sido especular que dentro del núcleo de cada célula del cuerpo hay una serie de instrucciones, escritas en grandes macromoléculas, en forma de hélices dobles apretadamente retorcidas en madejas para formar un grupo de cuarenta y seis cromosomas? El ADN de nuestro cuerpo desenrollado y engarzado llegaría hasta el Sol y volvería a la Tierra varias veces, diez o cien. Evidentemente, Darwin es el hombre que a conciencia descubrió una multitud de complejidades en las vidas y en los cuerpos de los percebes, las orquídeas y las lombrices y las describió con entusiasmo. Si en 1859 hubiera tenido un sueño profético acerca de las maravillas del ADN, sin duda lo habría revelado, aunque me pregunto si lo hubiera relatado con la seriedad en su semblante. Incluso para nosotros, que vivimos en la época de los ordenadores y estamos acostumbrados a los «milagros de la ingeniería», los hechos son difíciles de aceptar. No sólo máquinas copiadoras del tamaño de una molécula, sino enzimas que leen las pruebas y corrigen errores, todo ello a una velocidad deslumbrante, y a una escala que los superordenadores aún no pueden igualar. «Las macromoléculas biológicas poseen una capacidad de almacenamiento que excede la de los mejores depósitos de información actuales en varios órdenes de magnitud. Por ejemplo, la *densidad de información* en el genoma de la *Escherichia coli* es cercana a 10^{27} bits/m^3» (Küppers 1990:180).

En el capítulo 5 llegamos a una definición darwiniana de la posibilidad biológica en términos de accesibilidad dentro de la Biblioteca de Mendel, pero la precondición para esta Biblioteca, como hemos señalado, era la

existencia de mecanismos genéticos de asombrosa complejidad y eficiencia. William Paley habría sido transportado con admiración y curiosidad al intrincado nivel subatómico que hace posible la vida, ¿cómo podían haber evolucionado los mecanismos genéticos si son la precondición para la evolución darwiniana?

Los escépticos de la evolución argumentan que éste es el defecto fatal del darwinismo. Como hemos visto, el poder de la idea de Darwin proviene de la vía mediante la cual se distribuye la enorme tarea del diseño a través de las enormes cantidades de tiempo y espacio, preservando los productos parciales cuando procede. En el libro *Evolution: A Theory in Crisis*, Michael Denton lo expresa de esta forma: «Los darwinianos asumen que la existencia de islas de función es frecuente, que se encuentra con facilidad en el primer lugar, y que es fácil ir de isla en isla a través de intermediarios funcionales» (Denton 1985:317). Esto es casi correcto, pero no del todo. Naturalmente, la idea central de darwinismo es que el árbol de la vida extiende sus ramas, conectando «islas de función» con istmos de casos intermedios, pero nadie dice que el paso será «fácil» o seguros los lugares de parada «comunes». Sólo existe un sentido forzado de «facilidad» en el que el darwinismo está comprometido, de modo que estos «istmos que se cruzan» son fáciles de atravesar: dado que todo ser vivo es el descendiente de otro ser vivo, esto supone una tremenda ayuda; todo, incluso la más pequeña fracción de su receta, tiene la garantía de tener una viabilidad comprobada a través del tiempo. De acuerdo con el darwinismo, las líneas genealógicas son líneas de la vida. La única esperanza de entrar en este cósmico laberinto de material de desecho y seguir vivo es permanecer sobre los istmos.

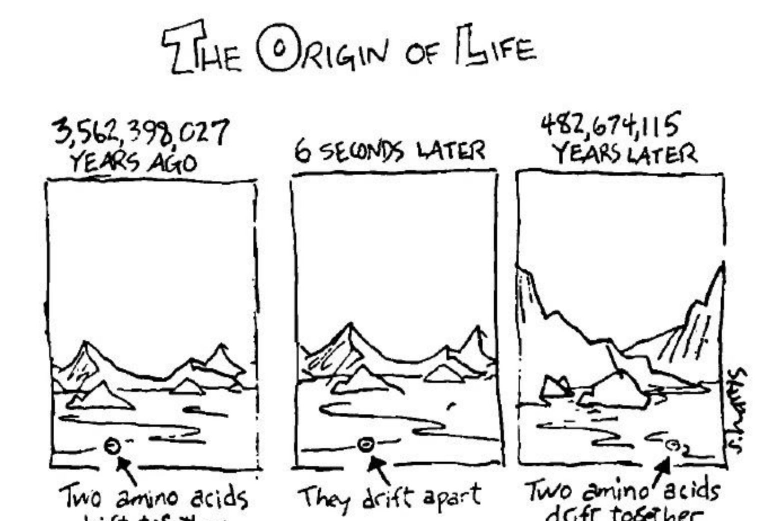


Figura 7.1

Pero ¿cómo puede haberse iniciado este proceso? Denton trata de calcular la improbabilidad de este inicio y llega a un número apropiado que provoca un aturdimiento de la mente.

Para conseguir al azar una célula se requeriría, al menos, que cien proteínas funcionales aparecieran simultáneamente en un lugar. Esto es cien sucesos simultáneos, cada uno de ellos con una probabilidad independiente, lo que podría ser más de 10^{-20} dada una máxima probabilidad combinada de 10^{-2000} (Denton 1985:323).

Esta probabilidad es pequeñísima, «evanescente», cercana a lo imposible. Y parece, a primera vista, como si la respuesta darwiniana estándar a tal reto no pudiera ayudarnos, desde un punto de vista lógico, puesto que las verdaderas condiciones previas para este éxito —un sistema de replicación con variaciones— son precisamente las únicas que permitirían explicarnos este éxito. Parece como si la teoría de la evolución se hubiera hundido en un profundo pozo del cual no puede salir. ¡Seguramente la única cosa que podría salvarla sería un «gancho celeste»! Esta era la ingenua esperanza de Asa Gray, y mientras más hemos aprendido acerca de lo intrincado que es el proceso de la replicación del ADN, más tentadora ha llegado a ser esta idea para aquellos que han estado buscando un lugar para sacar de apuros a la ciencia, con alguna ayuda para la religión. Se puede decir que para muchos esto es algo llovido del cielo. Olvidémoslo, dice Richard Dawkins:

Es posible, se argumenta, que el Creador no controle el día a día del proceso secuencia 1 en el que se inscriben los acontecimientos evolutivos, es posible que no haya concebido ni el tigre ni el cordero, es

posible que no hiciera un árbol, pero sí dispuso la maquinaria original de la replicación y del poder replicador, la maquinaria original del ADN y de las proteínas que realizan la selección acumulativa y que, en consecuencia, hacen posible toda la evolución.

Este es un argumento a todas luces débil, un argumento que, obviamente, cae por su propio peso. Es la complejidad organizada la que tenemos dificultad de explicar. Una vez que hemos aceptado *asumir* la complejidad organizada, aunque sólo sea la complejidad organizada del ADN/proteína, que funciona como un motor de replicación, es relativamente fácil invocar la existencia de un generador de todavía más complejidad organizada… Aunque, naturalmente, cualquier Dios capaz de diseñar inteligentemente algo tan complejo como la máquina de replicación ADN/proteína debe haber sido, al menos, tan complejo y organizado como la propia máquina (Dawkins 1986a:141).

Y Dawkins continúa: «Lo que hace de la teoría de la evolución una teoría bien ordenada es que explica cómo la complejidad organizada puede surgir de una simplicidad primitiva» (*ibidem*:316). Este es uno de los puntos fuertes de la idea de Darwin, y el punto débil de otras alternativas. En efecto, como ya he argumentado, es improbable que alguna otra teoría pueda poseer su fuerza:

Darwin describe un mundo de causas finales y leyes ideológicas con un principio que es, sin duda, mecanicista pero, lo que es más fundamental, absolutamente independiente de «significado» y de «propósito». Asume un mundo que es *absurdo* en el sentido existencialista del término: no ridículo sino sin sentido, y esta asunción es una condición necesaria de cualquier otra consideración del *propósito*, que no dé nada por probado. Es dudoso que podamos imaginar un principio no mecanicista pero que también no dé nada por probado para explicar el diseño en el mundo biológico; en este sentido, es tentador entender el compromiso con consideraciones que no den nada por probado, como equivalente a un compromiso con el materialismo mecanicista, aunque la prioridad de estos compromisos es clara… Alguien puede argumentar: la teoría materialista de Darwin puede no ser la única teoría que, en estos temas, no dé nada por probado, pero es esta teoría, sin duda, la única que hemos hallado como una buena razón para aceptar el materialismo (Dennett 1975:171-172).

¿Es ésta una crítica razonable e incluso apropiada de las alternativas religiosas? Un lector de uno de los primeros borradores de este capítulo se quejaba de este punto diciendo que al tratar la hipótesis de Dios precisamente como una más de las hipótesis científicas, para ser evaluada mediante los criterios de la ciencia en particular y del pensamiento racional en general, Dawkins y yo pasamos por alto la muy extendida opinión de los creyentes en Dios en el sentido de que su fe está más allá de la razón, y no es un asunto al que se le pueda aplicar los métodos habituales de testificación. No es que sea algo contrario a mí, aduce mi lector, sino totalmente indeseable, a mi entender, el asumir que el método científico continúa aplicándose con toda su potencia en este dominio de la fe.

De acuerdo, consideremos esta objeción. Dudo que a un defensor de la religión le parezca atractivo lo que voy a decir, una vez que lo analicemos cuidadosamente. El filósofo Ronald de Sousa definió en cierta ocasión, de manera memorable, la teología filosófica como «un partido de tenis intelectual que se juega sin red», y yo sin demora admito que he asumido, sin

comentario u objeción hasta ahora, que la red del juicio racional está levantada. Pero, si realmente mi lector lo desea, podemos bajarla. ¡Le toca servir!, le digo al defensor de la religión. Cualquiera que sea la «cosa» que sirva, supongo que le devolveré el servicio con rudeza, diciendo: «Lo que dice implica que Dios es un bocadillo de jamón envuelto en papel de aluminio. ¡Esto no es suficiente para que un Dios sea adorado!». Si me responde con una volea, queriendo conocer de qué forma puedo justificar lógicamente la afirmación de que su servicio tiene esa ridícula implicación con la posición de la red, le replicaré: «Entonces, ¿tal vez desea que la red esté levantada cuando le devuelva la pelota, y bajada en el momento de sus servicios? La red debe estar arriba o abajo. Si la red está bajada, entonces no hay reglas y cualquiera puede decir algo y —si hubiera alguno— participar en un juego MUG (siglas de Multi-User-Games, juegos de ordenador para muchos usuarios). Le he dado la posibilidad de asumir que no perdería su propio tiempo ni yo el mío jugando con la red bajada».

Ahora, si desea *razonar* acerca de la fe y ofrecer una razonable defensa (que responda a la razón) como una creencia de categoría extraordinaria, merecedora de una especial consideración, estoy deseando jugar. En realidad yo admito la existencia del fenómeno de la fe; lo que deseo es un fundamento razonado que me permita tomar la fe en serio, como una *vía para alcanzar la verdad*, y no, pienso, como una vía por la que la gente se conforta a sí misma y a los demás (una función meritoria, por otra parte, que acepto seriamente). Pero no debe esperar de mí que le siga en su defensa de la fe como un camino para alcanzar la verdad, si en cualquier momento apela a la dispensa («bajar la red») que está tratando de justificar. Antes de que apele a la fe, cuando la razón lo haya abandonado, que piense realmente si está dispuesto a abandonar la razón cuando ésta se encuentre de su lado. Imagínese que está haciendo turismo con una persona amada en un país extranjero y esta persona es brutalmente asesinada ante sus ojos. En el juicio resulta que en ese país los amigos del acusado pueden ser llamados como testigos por la defensa y testificar en favor de su inocencia. Observa el cortejo de sus compungidos amigos, obviamente sinceros, proclamando orgullosamente una fe inquebrantable en la inocencia del hombre al que vio cometer el horrible asesinato. El juez escucha asidua y respetuosamente, es obvio que más emocionado por aquella efusión que por toda la evidencia presentada por el fiscal. ¿No se trata de una pesadilla? ¿Viviría en ese país? O, poniendo otro ejemplo, ¿desearía ser operado por un cirujano que le dice que cada vez que oye en su interior una tenue voz que le dice que no tenga en cuenta lo

aprendido durante su formación como cirujano, él le hace caso? Sé que esto sucede en una compañía cortés, para permitir que la gente disponga de ambas opciones y, en la mayoría de las circunstancias, coopero de todo corazón con estas buenas disposiciones. Pero estamos tratando de llegar en serio a la verdad y si cree que esta común aunque innominable forma de entender la fe es algo mejor que la ofuscación socialmente útil para evitar situaciones embarazosas y de desprestigio, ha ahondado más en esta cuestión que cualquier filósofo (ninguno ha conseguido una buena defensa) o se está engañando a sí mismo (la pelota está ahora en su campo).

La réplica de Darwin al teórico que clamaría a Dios para dar el salto que inicie el proceso de la evolución es una refutación irrebatible, tan devastadora hoy como cuando Philo la utilizó dos siglos antes para vapulear a Cleantes en los *Diálogos* de Hume. Un «gancho celeste» representaría, cuando más, retrasar la solución del problema, pero Hume no podía pensar en grúas, así que claudicó. Darwin apareció con algunas grúas magníficas para llevar a cabo la elevación hasta un nivel *medio*, pero ¿pueden los principios que una vez funcionaron tan bien ser aplicados de nuevo para conseguir la elevación necesaria para levantar del suelo, en primer lugar, los pescantes de las grúas de Darwin? Sí. Justo cuando podría parecer que la idea darwiniana había llegado al final de sus recursos, salta elegantemente a un nivel *inferior* y mantiene en marcha no sólo una idea, sino muchas más que se multiplican como las escobas del aprendiz de brujo.

Si desea comprender dónde está el truco, que a primera vista parece inimaginable, tendrá que luchar con algunas ideas difíciles y un sinnúmero de detalles, tanto matemáticos como moleculares. Este no es el libro y yo no soy el autor, debe consultar aquellos detalles y no se puede asegurar realmente que los pueda entender, así que lo que sigue es una advertencia: aunque trataré de familiarizarlo con estas ideas, no las conocerá realmente si no las estudia en la literatura original. (Mi propia compresión de ellas es la de un aficionado). Muchos investigadores han puesto en marcha exploraciones imaginativas, teóricas y experimentales de las posibilidades, en lo que prácticamente constituye una subdisciplina situada en la frontera entre la biología y la física. Si no espero demostrarles la validez de estas ideas —y no deben confiar en mí si lo dijera—, ¿por qué las estoy presentando? Porque mi propósito es filosófico: deseo romper un prejuicio, la convicción de que un cierto tipo de teoría no podría *posiblemente* funcionar. Hemos visto cómo la trayectoria filosófica de Hume descendió por su incapacidad para tomar en serio una apertura en la pared que él, nebulosamente, vislumbró. Hume pensó

que sabía que no valía la pena proseguir en aquella dirección, y, como Sócrates, se cansó de señalarlo, pensar que se conoce cuando no se conoce es la causa principal de la parálisis filosófica. Si puedo demostrar que es concebible que la idea de Darwin puede transitar «hacia abajo por todo el camino» —hacia los orígenes de la vida—, esta demostración eliminará por adelantado una serie de rechazos por mero cumplimiento que nos son demasiado familiares, y abrirá nuestras mentes a otras posibilidades.

2. *Evolución molecular*

> Las moléculas de proteína más pequeñas de las células vivas, catalíticamente activas, consisten, al menos, en cien aminoácidos. Incluso para una molécula tan corta existen $20^{100} = 10^{130}$ disposiciones alternativas de los veinte monómeros básicos. Esto demuestra que incluso en los niveles más bajos de complejidad, el de las macromoléculas biológicas, es posible casi una ilimitada variedad de estructuras.
>
> Bernd-Olaf Küppers,
> *Information and the Origin of Life*

> Nuestra tarea es encontrar un algoritmo, una ley natural que nos conduzca al origen de la información.
>
> Manfred Eigen, *Steps towards Life*

Al describir, en la sección anterior, el poder de la idea central del darwinismo, me he ayudado con una pequeña (!) exageración: dije que todo ser vivo desciende de otro ser vivo. Esto no puede ser verdad, ya que implica la evidencia de una infinidad de seres vivos en un grupo sin ningún miembro que fuera el primero. Como sabemos que el número total de seres vivos (hasta ahora sobre la Tierra) es enorme pero finito, estamos obligados, lógicamente, a identificar al primer miembro: Adán o la protobacteria, sí se prefiere. Pero ¿cómo ha llegado a existir ese primer miembro? Una bacteria es, en su conjunto, demasiado complicada para que pueda llegar a existir como consecuencia de un accidente cósmico. El ADN de una bacteria, como la *E. coli* posee alrededor de cuatro millones de nucleótidos, casi todos ellos en un orden preciso. Además, es evidente que una bacteria no puede obtenerse con bastante menos. Así que nos encontramos con un dilema: dado que los seres vivos han existido sólo durante un tiempo finito, debe haber existido uno que fuera el primero, pero comoquiera que todos los seres vivos son complejos, ¡no podría haber existido uno que fuera el primero!

Sólo cabría una solución, y la conocemos bien en su planteamiento general: antes de que hubiera bacterias, con su metabolismo autónomo, habría habido otros seres bastante más simples, seres casi vivos como los virus, aunque, a diferencia de éstos, sin disponer todavía de la existencia de seres vivos en los que vivir como parásitos. Desde el punto de vista químico, los virus son «precisamente» cristales grandes y complejos, aunque, gracias a su complejidad, no están inactivos sino que «hacen cosas». En especial, los virus se reproducen o se autorreplican, con variaciones. Un virus viaja con ligereza, sin cargar con una maquinaria metabólica, por lo que o bien encuentra ocasionalmente la energía y los materiales necesarios para su autorreplicación o su autorreparación, o sucumbe a la segunda ley de la termodinámica y desaparece. Ahora, las células vivas facilitan almacenes concentrados para los virus, y éstos han evolucionado para explotarlos, pero en los días primigenios los virus tenían que ir en busca de otras vías menos eficientes para obtener copias de sí mismos. Hoy no todos los virus usan el ADN en doble hélice; algunos hacen uso de un lenguaje ancestral, compuesto de ARN en una sola hélice (lenguaje que naturalmente todavía desempeña un papel en nuestro propio sistema reproductivo, como un papel intermediario de «mensajero» durante la «expresión»). Si seguimos la práctica habitual y reservamos el término *virus* para una macromolécula parásita, necesitamos un nombre para los primeros ancestros. Los programadores de ordenadores denominan un «macro» a un fragmento de instrucciones codificadas —a modo de un empedrado— que realiza una particular función, así que propongo llamar *macros* a estos pioneros de la vida para resaltar que mientras son realmente grandes macromoléculas, son también «bits» de *programas* o *algoritmos*, mecanismos escuetos, mínimos, de autorreplicación, de modo muy parecido a los virus de los ordenadores que han emergido recientemente como una plaga para nuestra fascinación y preocupación (Ray 1992, Dawkins 1993)[51]. En cuanto estos pioneros macros se reproducen, cumplen las condiciones darwinianas necesarias para la evolución, y ahora está claro que emplearon la mejor parte de mil millones de años evolucionando sobre la Tierra, antes de que allí hubiera seres vivos.

Sin embargo, aun el más sencillo *macro* replicante está lejos de ser simple, ya que se trata de una composición de miles o millones de partes, según cómo contemos las materias primas de las que se componen. Las letras del alfabeto *A*denina, *C*itosina, *G*uanina, *T*imina y *U*racilo son bases no demasiado complejas como para que surgieran en el curso normal de los hechos prebióticos. (El ARN, que apareció antes que el ADN, contiene

*U*racilo, mientras que el ADN contiene *T*imina). Sin embargo, la opinión de los expertos difiere respecto a si estos bloques se pueden sintetizar por sí mismos mediante una serie de coincidencias precisamente en algo tan imaginativo como es un autorreplicador. El químico Graham Cairns-Smith (1982, 1985) ha presentado una versión actualizada del argumento de Paley, orientado al nivel molecular: el proceso de síntesis de fragmentos de ADN, incluso mediante los métodos más avanzados de la moderna química orgánica, es extraordinariamente complicado; esto demuestra que su creación por azar es tan improbable como el reloj de Paley en un vendaval. «Los nucleótidos son demasiado costosos» (Cairns-Smith 1985:45-49). El ADN requiere demasiado trabajo de diseño para ser un mero producto del azar, argumenta Cairns-Smith, pero entonces deduce una ingeniosa —aunque especulativa y controvertida— tesis sobre cómo podría haber sido realizado este trabajo. Sea o no sea eventualmente confirmada, la teoría de Cairns-Smith merece ser compartida ya que ejemplifica perfectamente la estrategia fundamental de Darwin[52].

Un buen darwiniano que se enfrenta de nuevo con el problema de encontrar una aguja en el pajar del espacio de diseño, buscaría por todas partes una forma *aún más simple* de replicador que pudiera servir como andamiaje temporal para mantener los componentes de la proteína o de los nucleótidos en su lugar, hasta que la totalidad de la proteína o los macros consiguieran conjuntarse. ¡Qué maravilloso sería poder decir que hay un candidato, precisamente con las propiedades adecuadas y, más maravilloso todavía, que es justamente lo que la Biblia ordenaba: barro! Cairns-Smith demostró que, además de los cristales aurorreplicantes de ADN y ARN basados en el carbono, hay otros cristales autorreplicantes bastante más simples (los que él llama de *baja tecnología*, basados en silicona, y estos silicatos, así denominados, pueden ser en sí mismos el producto de un proceso evolutivo). Estos cristales forman las ultrafinas partículas del barro que transforma las fuertes corrientes y los turbulentos remolinos en ríos, cristales que individualmente difieren, de manera sutil, a nivel de la estructura molecular, y que deciden cuándo «siembran» los procesos de cristalización que consiguen su autorreplicación.

Cairns-Smith desarrolló complicados argumentos para demostrar cómo los fragmentos de proteína y de ARN, que serán naturalmente atraídos a las superficies de los cristales como si fueran una multitud de pulgas, pueden eventualmente llegar a ser utilizados por los cristales de silicatos como «herramientas» con las que continuar sus propios procesos de

autorreplicación. De acuerdo con esta hipótesis (que, como todas las ideas fértiles, tiene muchas variaciones vecinas, cualquiera de las cuales puede llegar a ser la ganadora), los bloques de construcción de la vida inician su carrera como tipos casi parásitos, uniéndose para replicar partículas de barro y creciendo en complejidad para promocionar las «necesidades» de las partículas de barro hasta que alcanzan un punto en el que pueden defenderse por sí mismas. Ningún «gancho celeste», ni siquiera una escalera que después pudiera ser retirada, como Wittgenstein dijo una vez en otro contexto, una vez que se hubiera ascendido por ella.

Sin embargo esto no puede estar cerca de la historia total, aunque fuera verdad. Supongamos que cortas cadenas autorreplicantes de ARN llegan a ser creadas por ese proceso de baja tecnología. Cairns-Smith llamó «genes desnudos» a estos replicadores totalmente autoimplicados, debido a que no existen para ninguna otra cosa excepto *para* su propia autorreplicación, que realizan sin ayuda externa. Aún tenemos un gran problema: ¿cómo se visten estos genes desnudos? ¿Cómo estos solipsistas autorreproductores llegan a formar proteínas específicas, las enzimas, mínimas maquinarias que construyen los grandes cuerpos que transmiten los genes de hoy, de generación en generación? El problema es peor aún para estas proteínas —las enzimas—, que no construyen cuerpos y que son necesarias para asistir en el propio proceso de la autorreplicación, una vez que una cadena de ARN o de ADN se alarga. Mientras las cadenas cortas de ARN pueden replicarse a sí mismas sin la ayuda de enzimas, las cadenas largas necesitan un séquito de ayudantes y *especificarlas* requiere una secuencia muy larga; demasiado larga para que pueda ser replicada con una fidelidad suficientemente elevada sin la presencia de esas enzimas. Parece que nos enfrentemos de nuevo a una paradoja y entremos en un círculo vicioso descrito sucintamente por John Maynard Smith: «Uno no puede tener una replicación exacta sin una longitud de ARN, es decir, 2.000 pares de bases, y uno no puede tener bastante ARN sin una replicación exacta» (Maynard Smith 1979:445).

Uno de los investigadores líderes en este período de la historia de la evolución es Manfred Eigen. En su elegante y conciso libro titulado *Steps towards Life* (1992) —un buen libro para que continuemos la exploración de estas ideas— demuestra cómo los *macros* construyen gradualmente lo que él llama la «caja de herramientas molecular» que utilizan las células vivas para recrearse a sí mismas a la vez que también construyen a su alrededor las diferentes estructuras que llegarán a ser, a su debido tiempo, las membranas protectoras de las primeras células procariotas. Este largo período de

evolución precelular no ha dejado huellas fósiles, pero sí ha dejado multitud de claves de su historia en los «textos» que han sido transmitidos a través de sus descendientes entre los que se incluyen, naturalmente, los virus que pululan hoy a nuestro alrededor. Estudiando los textos que sobreviven actualmente, las secuencias específicas de A, G, G, y T en el ADN de los organismos superiores y las secuencias de A, C, G, y U del ARN, los investigadores pueden deducir muchas cosas acerca de la identidad actual de los textos autorreplicantes más precoces, utilizando versiones refinadas de las mismas técnicas que usan los filólogos para reconstruir las palabras que Platón realmente escribió. ¡Algunas secuencias en nuestro propio ADN son verdaderamente antiguas, incluso pueden seguirse (mediante una traducción hacia atrás en el lenguaje primitivo del ARN) en secuencias que se formaron en los primeros días de la evolución de los macros!

Retrocedamos al tiempo en el que las bases de nucleótidos (A, C, G, T, y U) estuvieron ocasionalmente presentes, allí y aquí, en cantidades variables, posiblemente congregadas alrededor de algunos de los cristales del barro de Cairns-Smith. Los veinte aminoácidos diferentes, que son bloques para la construcción de todas las proteínas, se presentan también con alguna frecuencia bajo una amplia gama de condiciones no bióticas, así que pueden servirnos de ayuda. Por otra parte, Sidney Fox (Fox y Dose 1972) ha demostrado que los aminoácidos individuales pueden condensarse para formar «proteinoides», sustancias parecidas a las proteínas que poseen una actividad catalítica muy modesta (Eigen 1992:32). Este es un paso pequeño pero importante, dado que la actividad catalítica —la capacidad de facilitar una reacción química— es el talento fundamental de cualquier proteína.

Supongamos ahora que algunas de las bases llegan a emparejarse, C con G, y A con U, y forman secuencias complementarias de ARN más pequeñas —menos de cien pares de longitud— que pueden replicar, regularmente, sin la ayuda de enzimas. En términos de la Biblioteca de Babel tendríamos en ese momento una imprenta y un encuadernador, pero los libros serían demasiado breves para ser buenos excepto para hacer más, con muchos errores de impresión. Y estos libros no podrían ser libros *acerca* de algo. Nos podría parecer correcto volver atrás, adonde han comenzado; o incluso peor. Cuando descendemos hasta el nivel de los bloques de construcción molecular, nos enfrentamos con un problema de diseño que es más parecido a la construcción de un juguete de hojalata que al modelado gradual de una escultura de barro. Bajo las rígidas reglas de la física, o los átomos saltan hasta reunirse en modelos estables o no lo hacen.

Afortunadamente para nosotros —también, naturalmente, para todos los seres vivos—, esparcidos en el «vasto» espacio de posibles proteínas, hay construcciones de proteínas que —si se encuentran— permiten que la vida siga adelante. ¿Cómo pueden encontrarse? De alguna manera tenemos que encontrar esas proteínas junto a los cazadores de proteínas, los fragmentos de cadenas de nucleótidos autorreplicantes que eventualmente llegan a «especificarlas» en los macros que componen. Eígen demostró cómo el círculo perverso puede convertirse en amistoso si se expande en un «hiperciclo» con más de dos elementos (Eigen y Schuster 1977). Se trata de un concepto técnicamente difícil, pero la idea subyacente es bastante clara: imaginemos una circunstancia en la que fragmentos de A pueden potenciar la perspectiva de buenos fragmentos de B, el que, a su vez, promueve el bienestar de *bits* de C, el que, completando el círculo, permite la replicación de más fragmentos de A, y así sucesivamente, en una comunidad de elementos que se refuerzan mutuamente hasta que se alcanza el punto desde el que la totalidad del proceso puede despegar, creando medios ambientes que normalmente sirven para replicar cadenas más y más largas de material genético (Maynard Smith 1979, es una gran ayuda en la comprensión de la idea del «hiperciclo»; véase también Eigen 1983).

Sin embargo, aunque esto sea posible en principio, ¿cómo pudo comenzar? Si todas las posibles proteínas y todos los «textos» posibles de nucleótidos tuvieran verdaderamente la misma probabilidad, sería muy difícil entender cómo el proceso pudo seguir adelante. De algún modo, la insulsa mezcla de ingredientes con aspecto de confeti tiene que alcanzar alguna estructura impuesta sobre ella, concentrando unos pocos candidatos «propicios a tener éxito» y, por lo tanto, haciéndolos aún más propicios a tener éxito. Recordemos los torneos de lanzamiento de monedas a cara o cruz de los que hablamos en el capítulo 2. Alguien tiene que ganar, pero el ganador lo logra en virtud de ningún valor, simplemente en virtud de un accidente histórico. El ganador no es más grande, más fuerte o mejor que los otros participantes, pero, no obstante, es el ganador. Algo similar sucede en la evolución molecular prebiótica, con un giro darwiniano: los ganadores consiguen hacer copias extras de sí mismos para la siguiente ronda, de modo que, sin ninguna selección «por alguna causa» (como dicen cuando no rechazan potenciales candidatos a jurados), comienzan a emerger dinastías de verdaderas hazañas replicativas. Si empezamos con un surtido totalmente aleatorio de «competidores» surgidos de la asociación de fragmentos autorreplicantes, aunque éstos no sean inicialmente distinguibles en lo que

respecta a su gesta replicadora, los que ganan en las primeras rondas ocuparán la mayoría de los puestos en las rondas siguientes, llenando el espacio con rastros de textos muy similares (cortos) pero dejando aún «vastos» hipervolúmenes del espacio absolutamente vacíos e inaccesibles para siempre. Los hilos iniciales de la protovida pueden emerger antes de que haya cualquier diferencia en habilidad, llegando a ser la realidad en la que puede crecer el árbol de la vida, gracias a los torneos de habilidad. Como señaló Bernd-Olaf Küppers [1990:150], colega de Eigen, «la teoría predice que las estructuras biológicas existen, pero no *cuáles*»[53]. Esto es suficiente para construir multitud de sesgos en el espacio de probabilidades desde el principio.

De este modo, algunos de los macros posibles, inevitablemente, son más probables —más probable encontrarse por casualidad en el «vasto» espacio de posibilidades— que otros. ¿Por qué éstos? ¿Acaso son los más «idóneos»? No lo son en ningún sentido no trivial pero sí en el sentido tautológico de ser idénticos (o casi idénticos) a quienes fueran previamente «ganadores», los cuales tienden, a su vez, a ser casi idénticos a los «ganadores» aún más antiguos. (En las dimensiones gigantescas de la Biblioteca de Mendel, las secuencias que difieren en un solo *locus* se colocan en los estantes «cercanas unas a otras» en alguna dimensión; la distancia de un volumen a otro es técnicamente conocida como la «distancia Hamming». Este proceso va separando gradualmente los «ganadores» —dando saltos de pequeñas «distancias Hamming»— desde cualquier punto de partida inicial en cualquier dirección de la Biblioteca). Este es el ejemplo posiblemente más rudimentario de la máxima «el rico se hace más rico», y dado que el éxito de una cinta tiene una explicación sin ninguna referencia más allá de la propia cinta y su semejanza con la cinta que la generó, ésta es una definición puramente sintáctica de adaptabilidad o idoneidad, opuesta a su definición semántica (Küppers 1990:141). Es decir, no es necesario tener en cuenta qué significa la cinta para determinar su «idoneidad». En el capítulo 6 hemos comprobado que simples cambios tipográficos no podrían nunca explicar el diseño que necesitamos explicar aunque podamos diferenciar la calidad entre dos libros mediante la comparación de sus frecuencias relativas en caracteres alfabéticos; pero antes de que podamos disponer del significado de los códigos de autorreplicación que haga esto posible, debemos tener los códigos de la autorreplicación que no significan «nada»; su «única» función es replicarse a sí mismos. Como Eigen ha señalado: «La estabilidad estructural de la molécula no se fundamenta en la información semántica que transporta,

la cual no se expresa hasta que aparece el producto de la traslación» (Eigen 1992:15).

Aquí está el nacimiento del fenómeno fundamental de QWERTY, pero, como el caso cultural que le da su nombre, no era, desde el comienzo, algo sin interés. La perfecta probabilidad similar podía haber sido disuelta en un monopolio por un proceso puramente aleatorio, como ya hemos visto, pero la perfecta probabilidad similar es dura de conseguir en la naturaleza en cualquier caso y en los mismísimos comienzos de este proceso de generación de textos, ya existía una predisposición. De las cuatro bases —A, C, G y T— G y C son las más estables estructuralmente: «El cálculo de la energía necesaria para los enlaces, a lo largo de experimentos de enlaces y síntesis, muestra que las secuencias ricas en G y C son mejores en la autorreplicación por instrucciones en plantilla sin la ayuda de enzimas» (Eigen 1992:34). Es decir, puede afirmarse que hay una predisposición, un sesgo en el deletreo, natural o físico. En inglés «e» y «t» aparecen más frecuentemente que «u» y «j» pero no porque las «es» y las «tes» sean más difíciles de borrar o más fáciles de fotocopiar o de escribir. (De hecho, la explicación marcha por otra vía; tenemos tendencia a utilizar los símbolos más fáciles de leer y escribir con las letras más frecuentemente usadas; por ejemplo, en el código Morse a la «e» se le asigna un solo punto y a la «t» una sola raya). En el ARN y el ADN esta explicación se invierte: G y C resultan favorecidas debido a que son las más estables en la replicación, no porque ocurran más frecuentemente en las «palabras» genéticas. Esta predisposición en el deletreo es precisamente «sintáctica» al principio, pero se combina con una predisposición *semántica*:

El examen del código genético (utilizando el llamado método «filológico») indica que sus primeros codones son muy ricos en G y en C. La secuencias GGC y GCC codifican respectivamente los aminoácidos glicina y alanina, y debido a su simplicidad química se forman en abundancia [en el mundo prebiótico]. La afirmación de que las primeras palabras del código fueron *asignadas* a los aminoácidos más comunes es de lo más plausible y subraya el hecho de que la lógica del esquema de la codificación resulta de leyes físicas y químicas y sus producciones en la naturaleza (Eigen 1992:34). (La cursiva es mía).

Estas «producciones» son *tipos diferentes de procesos algorítmicos* que incluyen las probabilidades o predisposiciones que se deben a las leyes de la física fundamentales y producen estructuras que, de otra forma, serían totalmente improbables. Como señala Eigen, el esquema resultante tiene una lógica; no se trata de dos cosas que llegan a unirse sino de una «tarea asignada», un sistema que llega a tener sentido y que tiene sentido debido a que —y sólo por eso— *funciona*.

Estos vínculos, en un principio «semánticos», son tan simples y locales que ya difícilmente cuentan como semánticos, aunque sin embargo pueden considerarse como una tenue luz de referencia; existe un enlace fortuito entre un *bit* de una cadena de un nucleótido con un fragmento de proteína que ayuda directa o indirectamente a reproducirlo. El círculo se cierra; una vez que este sistema de asignación semántico está en su lugar, todo se acelera. En este momento, un fragmento de la secuencia con el código puede ser el código *para* algo: una proteína. Esto crea una nueva dimensión de evaluación debido a que algunas proteínas son mejores que otras y hacen un trabajo catalítico y ayudan especialmente al proceso de replicación.

Todo esto abre las fronteras. Mientras que al principio las secuencias de macros pueden diferir solamente en su propia capacidad para autorreplicarse, ahora pueden ampliar sus diferencias creando otras estructuras más largas, y ligando su suerte a ellas. Una vez que se ha creado el asa del *feedback*, se pone en marcha una «carrera de armamentos»: macros cada vez más largos compiten por los bloques de construcción disponibles para construir sistemas autorreplicantes aún más grandes, más rápidos y más efectivos; aunque también más costosos. Nuestros sosos torneos de lanzamiento de monedas a cara o cruz, totalmente dependientes del azar, se han transformado a sí mismos en un torneo de habilidades. Desde este punto de vista, esto es mejor para la sucesión de ganadores que cuando se ganaba, trivialmente, lanzando la moneda.

¡Y el nuevo torneo funciona! Entre las proteínas hay muchas diferencias de «habilidades», de modo que existen amplias posibilidades de mejora, más allá de los minúsculos talentos catalíticos de los proteinoides. «En muchos casos, la catálisis enzimática acelera una reacción multiplicándola por un factor entre un millón y mil millones. Cada vez que este mecanismo se ha analizado cuantitativamente, el resultado ha sido el mismo: las enzimas son óptimos catalizadores» (Eigen 1992:22). El trabajo de catálisis realizado crea nuevos trabajos que han de ser cumplidos, así que el circuito de *feedback* se extiende para englobar oportunidades para la mejora más elaboradas. «Sea cual sea la tarea a la que se adapta una célula, ésta la cumple con óptima eficiencia. El alga azul-verde, un producto muy precoz de la evolución, transforma la luz en energía química con una eficiencia que se aproxima a la perfección» (Eigen 1992:16). Esta optimización no puede ser casual sino que debe ser el resultado de un proceso gradual autodirigido de mejoramiento. De este modo, de un grupo de pequeñas predisposiciones en las probabilidades y

competencias iniciales de los bloques de construcción, un proceso de automejora, como una bola de nieve, se ha puesto en marcha.

3. *Las leyes del juego de la vida*

> Este bellísimo sistema formado por el sol, los planetas y los cometas, sólo puede proceder del consejo y del dominio de una Inteligencia y de un Ser Poderoso.
>
> Isaac Newton
> (Citado en Ellegard, «*The Darwinian Theory and the Argument from Design*»)

> Cuanto más examino el universo y más estudio los detalles de su arquitectura, más evidencias encuentro de que este universo, en algún sentido, debe haber sabido que estábamos llegando.
>
> Freeman Dyson, *Disturbing the Universe*

> Es fácil imaginar un mundo que, aunque ordenado, no posea, sin embargo, el correcto tipo de fuerzas o de condiciones para su salida desde una profundidad significativa.
>
> Paul Davies, *La mente de Dios*

Afortunadamente para nosotros, las leyes de la física aceptan que en el «vasto» espacio de las posibles proteínas haya macromoléculas de una virtuosidad catalítica tan impresionante que puedan servir como bloques de construcción activos de la complejidad de la vida. Y, con la misma fortuna, las mismas leyes de la física facilitan la existencia en el mundo del suficiente desequilibrio, de modo que los propios procesos algorítmicos puedan iniciar un salto que conduzca eventualmente al descubrimiento de las macromoléculas, convirtiéndolas en herramientas para otra oleada de exploraciones y descubrimientos. ¡Demos gracias a Dios por la existencia de estas leyes!

¿Razonable? ¿O no deberíamos considerarlo así? Como hemos visto, si las leyes fueran diferentes, el árbol de la vida podía no haber brotado nunca. Podemos tratar de excusar a Dios de la tarea de diseñar el sistema de la replicación con su maquinaria (la cual se puede diseñar a sí misma automáticamente, si algunas de la teorías discutidas en la sección anterior de este libro son correctas o se mueven en el camino adecuado), pero incluso si admitimos que esto es así, aún contamos con el hecho de que las leyes físicas

permiten que este admirable despliegue suceda, lo que ha sido suficiente para que mucha gente se pregunte si la inteligencia del creador es la sabiduría del legislador, en lugar de la inventiva del ingeniero.

Cuando Darwin acariciaba la idea de que las leyes de la naturaleza eran diseñadas por Dios se encontraba bien acompañado, en su pasado y en su presente. Newton insistió en que la original disposición original del universo era inexplicable por «simples causas naturales» y sólo se podía atribuir al «consejo y a la inventiva de un agente voluntario». Einstein se refería a las leyes de la naturaleza como «los secretos del Anciano» y es famoso cómo expresaba su incredulidad sobre el papel del azar en la mecánica cuántica proclamando que «Dios no juega a los dados». Más recientemente, el astrónomo Fred Hoyle ha dicho: «No creo que cualquier científico que examinase la evidencia dejase de aplicar la inferencia de que las leyes de la física nuclear han sido deliberadamente diseñadas con respecto a las consecuencias que producen en el interior de las estrellas» (citado en Barrow y Tipler, 1988:22). El físico y cosmólogo Freeman Dyson trató esta cuestión con mucha mayor cautela diciendo: «No sostengo que la arquitectura del universo pruebe la existencia de Dios. Sólo digo que la arquitectura del universo es consistente con la hipótesis de que la mente tiene un papel esencial en su funcionamiento» (Dyson 1979:251). El propio Darwin estaba preparado para conceder una honorable tregua a esta cuestión, aunque el pensamiento de los darwinianos seguía adelante con el fuerte impulso creado por el éxito de sus primeras aplicaciones de la misma cuestión en otros contextos.

Cuando se ha ido averiguando más sobre el desarrollo del universo desde el Big Bang, sobre las condiciones que han permitido el desarrollo de las galaxias y de las estrellas y de los elementos pesados de los que pueden haberse formado los planetas, los físicos y los cosmólogos se han impresionado cada vez más por la exquisita sensibilidad de las leyes de la naturaleza. La velocidad de la luz es aproximadamente de 186.000 millas por segundo. ¿Qué sucedería si esta velocidad fuera sólo de 185.000 millas por segundo o bien de 187.000 millas por segundo? ¿Provocaría esta modificación algunos cambios? ¿Qué pasaría si la fuerza de la gravedad fuera un 1 por ciento más elevada o menos elevada de lo que es? Las constantes fundamentales de la física —la velocidad de la luz, la constante de la atracción gravitatoria, las fuerzas débiles y fuertes de la interacción subatómica, la constante de Planck— presentan valores que naturalmente permiten el desarrollo real del universo tal como sabemos que ha sucedido.

Pero resulta que si con la imaginación modificamos algunos de estos valores en mínimas cantidades, tendríamos como consecuencia un universo en el que nada de esto habría sucedido y en el que, aparentemente, nada parecido a la vida podría haber surgido: ni planetas, ni atmósferas, ni sólidos, ni elementos, excepto oxígeno y helio, o quizás incluso ni éstos, sólo algún aburrido plasma de calor, materia indiferenciada, o una nada igualmente aburrida. ¿No es un hecho admirable que las leyes de la naturaleza sean exactamente las *adecuadas* para que existamos? ¡Naturalmente que uno puede caer en la tentación de añadir que nosotros no las fabricamos!

¿Este admirable hecho necesita una explicación y, si es así, qué tipo de explicación puede recibir? De acuerdo con el principio antrópico, estamos autorizados para inferir hechos acerca del universo y sus leyes a partir del hecho indiscutible de que nosotros (*anthropoi*, es decir, seres humanos) estamos aquí para inferir y observar. El principio antrópico se presenta con varias formas. (Entre los libros recientes que son de utilidad están los de Barrow y Tipler 1988, y Breuer 1991. Consúltense también los libros de Pagels 1985 y de Gardner 1986).

En su «forma débil», el principio antrópico es una aplicación de la lógica fundamental bien fundada, inocua *y*, en ocasiones, útil: si *x* es una condición necesaria para la existencia de *y* y existe y, entonces *x* existe. Si la conciencia depende de estructuras físicas complejas y las estructuras físicas complejas dependen de grandes moléculas compuestas por elementos más pesados que el hidrógeno y el helio, en estas circunstancias, dado que somos conscientes, el mundo debe contener tales elementos.

En realidad, en el párrafo anterior hay una bomba de relojería encubierta: la imprecisa expresión *debe*. He seguido la práctica común de aplicar una expresión de necesidad (*debe*) de una manera técnicamente incorrecta. Como cualquier estudiante de lógica aprende enseguida, lo que realmente debería haber escrito es: «*Se podría suponer que*: Si la conciencia depende…, dado que somos conscientes, el mundo *contiene* tales elementos».

La única conclusión válida que podemos extraer es que el mundo *contiene* tales elementos, y no que tiene que contener tales elementos. Podemos admitir que el universo *tiene* que contener tales elementos *para que nosotros existamos*, pero puede no haber contenido tales elementos, y si este hubiera sido el caso, no estaríamos aquí para afligirnos. Tan simple como esto.

Algunos intentos para definir y defender una «variante fuerte» del principio antrópico se esfuerzan en justificar la última localización del *debe* no como una expresión casual, sino como una conclusión acerca de la vía por

la que el universo necesariamente es. Admito que me ha parecido difícil creer que tanta confusión y controversia hayan sido generadas realmente por un simple error de lógica, pero la evidencia es muy fuerte en favor de que se trata de esto y no precisamente de discusiones sobre el principio antrópico. Consideremos las confusiones que rodean a la deducción darwiniana en general. Darwin deduce que los seres humanos *deben haber* evolucionado a partir de un ancestro común con el chimpancé, o que toda la vida *debe haber* surgido de un único comienzo, y algunos, irresponsablemente, toman estas deducciones como afirmaciones de que los seres humanos son de algún modo un producto necesario de la evolución, o bien que la vida es un hecho necesario de nuestro planeta. Aunque si las deducciones de Darwin se construyen apropiadamente, no se puede deducir nada parecido. La cuestión no debe ser que nosotros estemos aquí, sino que *dado* que estamos aquí, hemos evolucionado desde los primates. Supongamos que Juan es soltero. Entonces, *debe* ser un hombre no casado. ¿Correcto? (Esto es una verdad lógica). ¡Pobre Juan, nunca conseguirá casarse! La falacia es obvia en este ejemplo y merece la pena que la guardemos en la memoria como un modelo que nos permitirá su comparación con otros argumentos.

Los creyentes en alguna de las «versiones fuertes» propuestas por el principio antrópico piensan que pueden deducir algo admirable y sorprendente del hecho de que nosotros, observadores conscientes, estemos aquí. Por ejemplo, que en algún sentido el universo existe *para* nosotros, o quizá que nosotros existimos *porque* el universo, como globalidad, existe, o incluso que Dios creó el universo para que nosotros existiéramos. Construidas de esta forma, estas propuestas son intentos de restaurar el argumento del diseño de Paley, redirigiéndolo al diseño de las leyes de la física del universo más generales, no a las construcciones particulares que dichas leyes han hecho posibles. Aquí, una vez más, hay disponibles contraofensivas darwinianas.

Estas son aguas profundas y la mayoría de las discusiones de los temas se entremezclan con términos técnicos, pero la fuerza lógica de estas respuestas darwinianas puede ser evocada brillantemente si se considera un caso bastante más simple. En primer lugar, debo introducir el juego de la vida, un excelente *meme* cuyo principal autor es el matemático John Horton Conway. (Utilizaré esta valiosa herramienta mental con otros objetivos más adelante. Este juego realiza un excelente trabajo en el planteamiento de un tema complicado al reflejar únicamente su simple esencia o esqueleto, lista para ser comprendida y valorada).

La vida se juega en una cuadrícula bidimensional, como un tablero de ajedrez, utilizando unas simples fichas, que pueden ser cantos rodados o monedas, aunque también puede optarse por la alta tecnología y jugar en la pantalla de un ordenador. No se trata de un juego competitivo, sino de un solitario[54]. La cuadrícula divide el espacio en casillas cuadradas, y cada casilla se encuentra, en cada momento, activada (*on*) o desactivada (*off*). (Si la casilla se encuentra en posición *on* —es decir, activada—, hay que colocar una moneda; si se encuentra en posición *off* —o sea, desactivada—, hay que dejarla vacía). Obsérvese que en la figura 7.2 cada casilla tiene a su alrededor ocho casillas vecinas: las cuatro casillas contiguas —norte, sur, este y oeste— y las cuatro diagonales —nordeste, sureste, suroeste y noroeste.

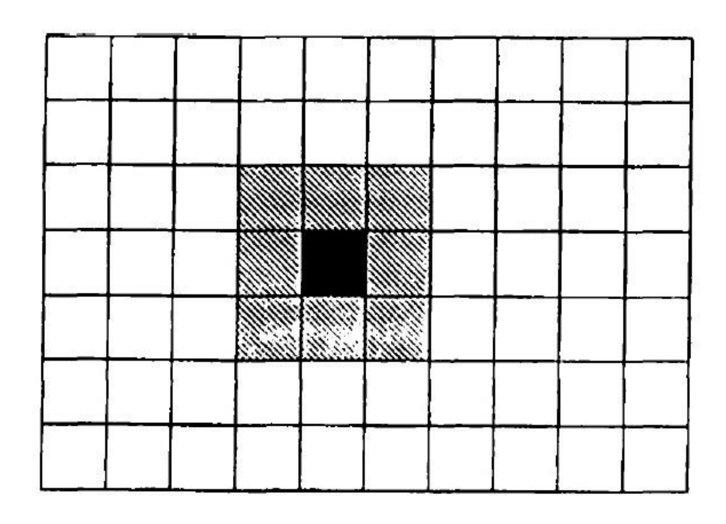


Figura 7.2

El tiempo en el mundo de la vida es discreto, discontinuo; avanza en latidos, y el estado del mundo cambia entre cada dos latidos de acuerdo con la siguiente regla:

Física de la vida: para cada casilla de la cuadrícula, cuéntese cuántas de sus ocho casillas vecinas hay activadas (*on*) en este instante. Si la respuesta es exactamente dos, la casilla permanece en su estado actual (*on* o *off*) en el siguiente instante. Si la respuesta es exactamente tres, la casilla es activada (*on*) en el siguiente instante, cualquiera que sea su estado actual. Bajo todas las otras condiciones la casilla se encuentra desactivada (*off*).

Esta es la única regla del juego. Ahora sabemos todo lo que hay que saber para jugar al juego de la vida. *La totalidad de la física del mundo de la vida se encuentra incluida en esta única y nada excepcional ley*. En tanto que es la ley fundamental de la «física» del mundo de la vida, nos ayuda, en primer lugar, a concebir esta curiosa física en términos biológicos: pensemos que las

casillas que se activan (*on*) son nacimientos, las casillas desactivadas (*off*) son muertes y los sucesivos instantes constituyen generaciones. Tanto la superpoblación (más de tres casillas vecinas habitadas) como el aislamiento (menos de dos casillas vecinas habitadas) conducen a la muerte. Veamos unos cuantos casos simples.

En la configuración de la figura 7.3 sólo las casillas *d* y *f* tienen cada una tres vecinos en posición *on*, así que éstas serán las únicas casillas nacidas en la siguiente generación.

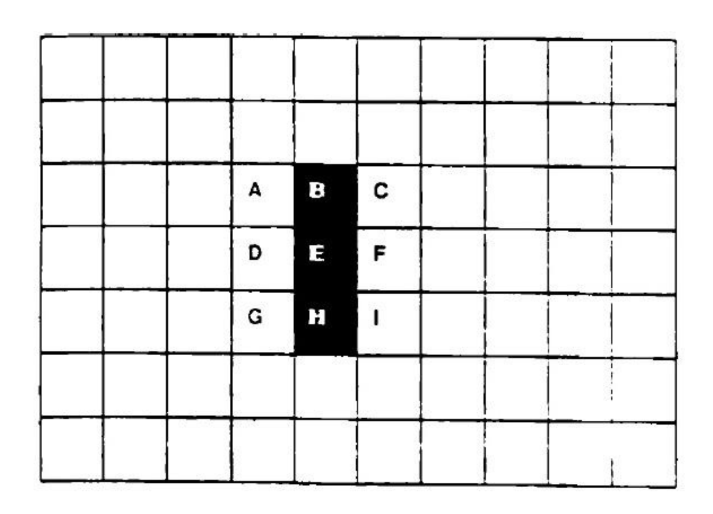


Figura 7.3

Las casillas *b* y *h* tienen cada una sólo un vecino en *on*, así que ellas mueren en la siguiente generación. La casilla *e* tiene dos vecinos en *on*, así que permanece activada. De este modo el siguiente «instante» será la configuración que se muestra en la figura 7.4.

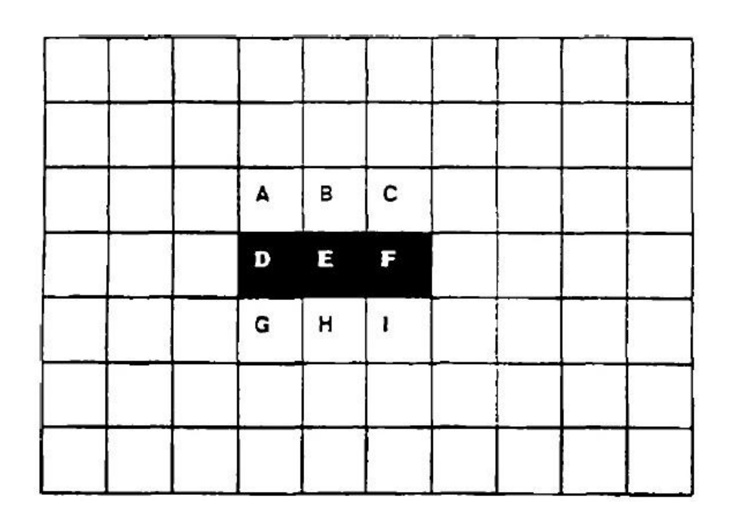


Figura 7.4

Obviamente, la configuración revertirá en el siguiente instante, y este pequeño modelo cambiará de forma hacia delante y hacia atrás indefinidamente, a menos que algunas nuevas casillas activadas [en *on*] aparezcan en el dibujo. Esto es lo que se llama destello o luz de tráfico. ¿Qué sucederá en la configuración de la figura 7.5?

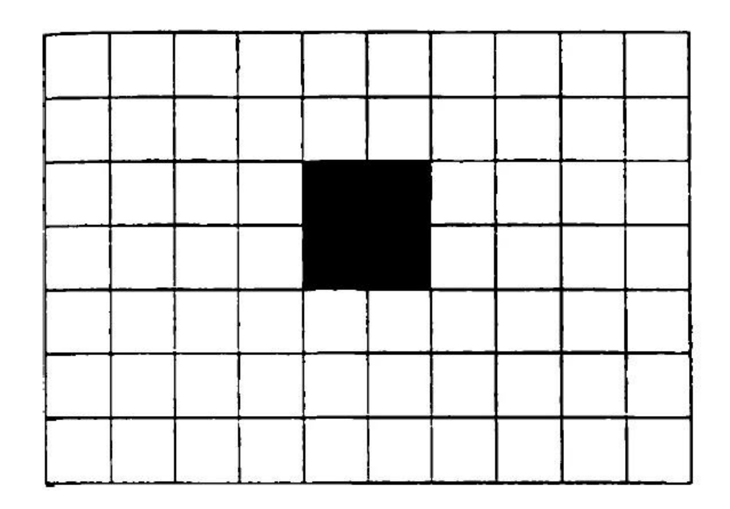


Figura 7.5

Nada. Cada casilla en situación activada [*on*] tiene tres celdillas vecinas en *on*, así que no se produce ningún nacimiento. Esta configuración se denomina «naturaleza muerta». Mediante la escrupulosa aplicación de nuestra simple ley se puede predecir con perfecta precisión el próximo instante de cualquier configuración de casillas en *on* y en *off*, y el instante después de éste y así sucesivamente. En otras palabras, el mundo de la vida es un mundo-

juguete que ejemplifica perfectamente el determinismo que Laplace hizo famoso: si se nos da la descripción del estado de este mundo en un instante, nosotros, como observadores, podemos predecir con precisión los futuros instantes mediante la simple aplicación de nuestra ley física. O, dicho en términos que ya he utilizado en escritos anteriores (Dennett 1971, 1978, 1987b), *cuando adoptamos la posición física* hacia una configuración del mundo de la vida, nuestros poderes de predicción son perfectos: no hay ruido, no hay incertidumbre, ninguna probabilidad es menor que otra. Por otra parte, como consecuencia del carácter bidimensional del mundo de la vida, nada está oculto a la vista. No hay trastienda; no hay variables ocultas; el despliegue de la física de los objetos en el mundo de la vida es directa y completamente visible.

Si opina que el seguimiento de esta regla es un ejercicio aburrido, tiene la posibilidad de utilizar simulaciones del mundo de la vida en el ordenador en las que se pueden establecer configuraciones en la pantalla y dejar que el ordenador ejecute el algoritmo, cambiando la configuración una y otra vez de acuerdo con la regla. En las mejores simulaciones, se puede cambiar la escala tanto del tiempo como del espacio, alternando entre una aproximación con primer plano y una vista de pájaro. Un toque elegante añadido a algunas versiones en color es que las casillas en *on* (a menudo denominadas *pixels*) están codificadas en colores de acuerdo con su edad: nacen de color azul y cambian de color en cada generación, pasando por los colores verde, amarillo, naranja, rojo, marrón y negro, color en el que permanecen hasta que mueren. Esto permite darse cuenta con una sola mirada de lo viejos que son algunos modelos, qué casillas son de la misma generación, dónde se producen nacimientos y muchas cosas más[55].

Pronto se descubre que algunas configuraciones simples son más interesantes que otras. Contemplemos un segmento en diagonal como el de la figura 7.6.

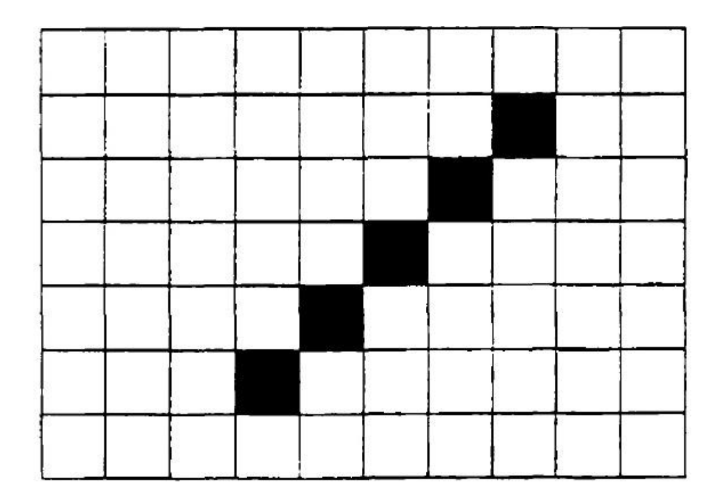


Figura 7.6

Éste *no* es un destello; en cada generación, sus dos casillas en *on* de los extremos mueren por aislamiento y no nacen otras. Todo el segmento se evapora con rapidez. Además de la configuración que nunca cambia —la «naturaleza muerta»— y aquellas que se evaporan rápidamente —como la del segmento en línea diagonal—, hay configuraciones con todo tipo de periodicidad. El destello, como vemos, tiene un período de dos generaciones que continúa *ad infinitum*, a menos que otra configuración la usurpe. Esta usurpación es lo que hace interesante a la vida: entre las configuraciones periódicas hay algunas que se deslizan, como una ameba, a través del plano. La más simple es el «deslizador», una configuración de 5 celdillas coloreadas [*pixels*] que se muestra en la figura 7.7, en un simple movimiento hacia el sureste.

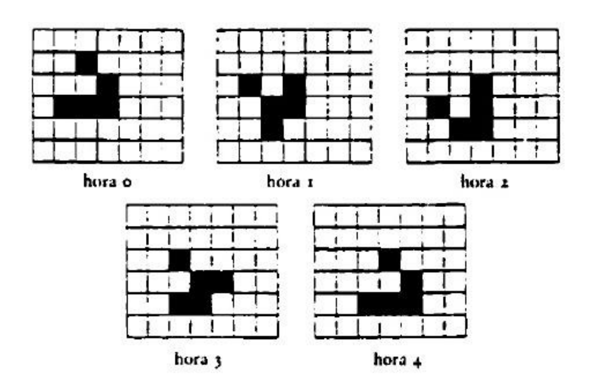


Figura 7.7

También están las configuraciones conocidas como los «comilones», las «locomotoras de vapor», los «rastreadores del espacio»; y muchas otras denominadas «ciudadanos» del mundo de la vida que emergen como objetos reconocibles a un nuevo nivel. (Este nivel es análogo al que en un trabajo previo hemos denominado el *nivel del diseño*). Este nivel tiene su propio lenguaje, un transparente escorzo de las aburridas descripciones que uno daría al nivel físico. Por ejemplo:

Un «comilón» puede comerse un «deslizador» en cuatro generaciones. Cualquiera que se coma, el proceso básico es el mismo. Se forma un puente entre el «comilón» y su presa. En la siguiente generación la región del puente muere por superpoblación, incluyendo un bocado del propio «comilón» y otro de la «presa». El «comilón» se repara entonces a sí mismo. La presa generalmente no puede hacerlo. Si el resto de la presa se extingue, como el «deslizador», la presa se consume (Poundstone 1985:38).

Llama la atención algo curioso que sucede a nuestra «ontología» —nuestro catálogo de lo que existe— cuando nos movemos entre niveles. En el nivel físico no hay movimiento, solamente *on* y *off*, y la única cosa individual que existe, la casilla, se define por su posición espacial fija. En el nivel del diseño tenemos súbitamente la movilidad de objetos persistentes; es uno y el mismo «deslizador» (aunque en cada generación compuesto por diferentes casillas) el que se ha movido hacia el sureste en la figura 7.6, cambiando su forma cuando se mueve; y existe un «deslizador» menos en el mundo después que el «comilón» se lo haya comido en la figura 7.8.

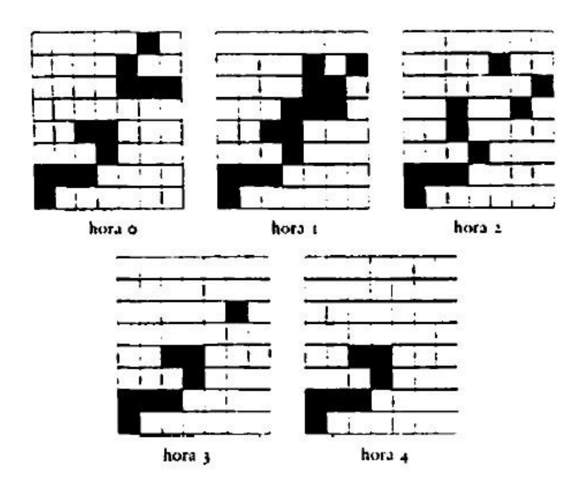


Figura 7.8

Obsérvese, también, que mientras que en el nivel físico no se encuentran excepciones a la ley general, en este nivel nuestras generalizaciones deben restringirse: requieren la introducción de cláusulas tales como «usualmente» o

«a condición de que nada las usurpe». En este nivel, *bits* descarriados de basura de anteriores acontecimientos pueden «romper» o «matar» uno de los objetos de la ontología. Su *prominencia como cosas reales* es considerable, pero no está garantizada. Al decir que su prominencia es considerable se expresa que uno puede, con algún pequeño riesgo, ascender a este nivel de diseño, adoptar su ontología y proceder a predecir —a grandes rasgos y arriesgadamente— la conducta de las configuraciones más grandes o sistemas de configuraciones sin preocuparse de computar el nivel físico. Por ejemplo, uno puede asignarse la tarea de diseñar algún supersistema interesante a partir de las «partes» que el nivel del diseño hace disponibles.

Esto es justamente lo que Conway y sus estudiantes se dispusieron a hacer y el éxito fue majestuoso. Diseñaron y probaron la viabilidad del diseño, una entidad autorreproductora compuesta enteramente de casillas de la vida que eran también (en gran medida) una máquina de Turing universal; ¡un ordenador de dos dimensiones que, en principio, puede computar cualquier función calculable! ¿Qué pudo ser lo que inspiró a Conway y a sus estudiantes a crear, en primer lugar, este mundo y después ese asombroso habitante de este mundo? Su intención era responder, a un nivel muy abstracto, a una de las cuestiones centrales que hemos estado considerando en este capítulo: ¿cuál es la mínima complejidad que se requiere para que se autorreproduzca una cosa? Estaban siguiendo las brillantes y pioneras especulaciones de John von Neumann, que había estado trabajando en esta cuestión hasta su muerte en el año 1957. Francis Crick y James Watson descubrieron el ADN en 195 3, aunque el mecanismo de su replicación fue un misterio durante muchos años. Von Neumann había imaginado, con algún detalle, una especie de robot flotante que recuperaba pecios y otros fragmentos flotantes para utilizarlos en la construcción de su duplicación, el cual sería capaz de repetir el proceso. Su descripción de cómo un autómata leía su propio proyecto y lo copiaba en una nueva creación (publicada postumamente en 1966) anticipaba con increíble detalle muchos descubrimientos posteriores acerca del mecanismo de la expresión del ADN y de su replicación, aunque con el fin de probar la posibilidad de un autómata matemáticamente autorreproducible, riguroso y tratable, Von Neumann había recurrido a una simple y bidimensional abstracción, ahora conocida como *autómata celular*. Las casillas del juego de la vida de Conway son un ejemplo muy aceptable del *autómata celular*.

El objetivo de Conway y sus estudiantes era confirmar la prueba de Von Neumann en detalle, aunque realmente construyeron un mundo bidimensional

con una física simple en el que tal construcción autorreplicante sería una estructura estable y *funcionante*. Como pretendía Neumann, deseaban que su respuesta fuera lo más general posible, y en consecuencia, lo más independiente posible de la física y química actual (¿terrestre?, ¿local?). Deseaban algo tremendamente simple, fácil de visualizar y fácil de calcular, de modo que no se limitaron a pasar de tres dimensiones a dos; también «digitalizaron» tanto el espacio como el tiempo; todos los tiempos y todas las distancias, como vemos, están expresados en números totales de «instantes» y «casillas». Fue Neumann quien utilizó la abstracta concepción de Alan Turing de un ordenador mecánico (ahora denominado «máquina de Turing») y lo construyó con la especificación de un ordenador, general en su propósito, que almacena programas y que procesa seriadamente (ahora llamado la «máquina de Neumann»); en su brillante exploración de los requerimientos espaciales y estructurales para tal ordenador, Conway había realizado —y probado— que una máquina universal de Turing (una máquina de Turing que puede computar totalmente cualquier función calculable) puede, en principio, ser construida en un mundo bidimensional[56]. Conway y sus estudiantes también consiguieron confirmarlo con su propio ejercicio de ingeniería en dos dimensiones[57].

No era nada fácil, pero demostraron que se podía construir un ordenador con las formas más simples de la vida. Las corrientes de «deslizadores» pueden proporcionar, por ejemplo, la «cinta» de *input-output* y el lector de la cinta puede ser un voluminoso conjunto de «comilones», «deslizadores» y otros *bits* y piezas. ¿Qué aspecto tendría esta máquina? Poundstone ha calculado que la construcción total tendría alrededor de 10^{13} casillas o *pixels*.

Desplegar un modelo de 10^{13} pixels requeriría una pantalla de vídeo que rondaría los 3 millones de *pixels*. Asumiendo que los pixels tuvieran 1 milímetro cuadrado (un estándar de resolución muy elevada para los ordenadores personales), la pantalla tendría 3 kilómetros, lo que supone un área seis veces la de Monaco.

La perspectiva de un modelo autorreproducible reduciría los *pixels* hasta la invisibilidad. Si nos alejamos de la pantalla para ver cómodamente la totalidad del modelo, los pixels (e incluso los «deslizadores», «comilones» y pistolas) serían demasiado pequeños. Un modelo que se autorreprodujera sería una masa nebulosa, como una galaxia (Poundstone 1985:227-228).

En otras palabras, cuando hubiésemos encajado bastantes piezas en algo que pudiera reproducirse a sí mismo (en un mundo bidimensional), el resultado sería muchísimo más grande que sus *bits* más pequeños, del mismo modo que un organismo es mucho más grande que sus átomos. Probablemente no podremos hacerlo con algo mucho menos complicado, aunque esto, en sentido estricto, no ha sido probado. El presentimiento con el que

comenzamos este capítulo consigue un apoyo magnífico: llegar a disponer de los *bits* y las piezas necesarios para construir un objeto autorreplicante exige *muchísimo* trabajo de diseño (el trabajo realizado por Conway y sus estudiantes); los autorreplicadores no coparticipan en coincidencias cósmicas; son demasiado grandes y costosos.

El Juego de la Vida ilustra muchos principios importantes y puede ser utilizado para construir diferentes argumentos o experimentos mentales, aunque en este momento me contentaría con dos cuestiones que son especialmente relevantes para esta fase de nuestro argumento antes de volver a la cuestión principal (para más reflexiones sobre el Juego de la Vida y sus implicaciones, véase Dennett 1991b).

En primer lugar, obsérvese que la distinción entre orden y diseño se ha difuminado, tal como le sucedió a Hume. Conway *diseñó* la totalidad del mundo de la vida, es decir, pretendía articular un orden que *funcionara* de una cierta manera. Pero ¿deben considerarse los «deslizadores», por ejemplo, como cosas diseñadas o como objetos naturales, igual que los átomos o las moléculas? Seguramente el *lector de la cinta* que Conway y sus alumnos dispusieron fuera de los «deslizadores» y similares, es un objeto diseñado, pero el más simple de los «deslizadores» parece que se encontraría «automáticamente» fuera de la física básica del mundo del Juego de la Vida; nadie diseñó o inventó el «deslizador»; tan sólo *se descubrió* su implicación en la física del mundo del Juego de la Vida. Pero esto que decimos es realmente verdad en *todas las cosas* en el mundo de la vida. Nada sucede en el mundo de la vida si no está estrictamente implicado —lógicamente deducible por demostración directa de teoremas— en la física y en la configuración inicial de las casillas. Algunas de las cosas en el mundo de la vida son realmente más maravillosas y más impredecibles (por nosotros, con nuestra inteligencia, propicia a los juegos de palabras) que otras. En cierto sentido, la galaxia de *pixels* de Conway que se autorreproduce es una más de las macromoléculas de la Vida con una muy larga y complicada periodicidad en su conducta.

¿Qué pasaría si pusiéramos en marcha una enorme multitud de estos autorreproductores, y les permitiéramos competir por recursos? Supongamos que evolucionan, es decir, que sus descendientes no son exactamente sus duplicados. ¿Serían estos descendientes más exigentes al haber sido diseñados? Quizás, aunque no se puede trazar una línea entre las cosas simplemente ordenadas y las cosas diseñadas. El ingeniero comienza con algunos *objets trouvés*, encuentra objetos con propiedades que pueden ser

aprovechados en construcciones más amplias, pero las diferencias entre un clavo diseñado y manufacturado, un tablón aserrado, y una plancha de pizarra que se encuentra en la naturaleza no son «principios fundamentales». Las alas de las gaviotas son magníficos elevadores, las moléculas de hemoglobina son soberbias máquinas transportadoras, las moléculas de glucosa excelentes paquetes de energía y los átomos de carbono excelentes pegamentos para todo tipo de uniones.

En segundo lugar, el Juego de la Vida es una excelente ilustración de la potencia —y de una concomitante debilidad— de las simulaciones en ordenador dirigidas a solucionar las cuestiones científicas. Suele suceder que la única vía para persuadirnos de generalizaciones muy abstractas era probarlas con rigor a partir de los principios fundamentales o axiomas de cualquier teoría: matemática, física, química, economía. A principios de este siglo, empezó a quedar claro que muchos de los cálculos teóricos que apetecería hacer en cualquiera de estas ciencias estaban sencillamente más allá de la capacidad humana, es decir, eran «intratables». Llegaron entonces los ordenadores y abrieron una nueva vía para enfrentarse con estas cuestiones: las simulaciones masivas. La simulación del tiempo es el ejemplo más conocido por todos nosotros cuando vemos en la televisión a los meteorólogos, pero la simulación con ordenadores también está revolucionando el trabajo científico en muchos otros campos, hasta convertirse, probablemente, en el más importante avance epistemológico del método científico desde la invención de instrumentos más precisos para la medida del tiempo. En la teoría de la evolución, ha nacido recientemente la nueva disciplina de la Vida Artificial que facilita una denominación y un paraguas que cubre a una verdadera carrera, una carrera del oro, a la que se han lanzado investigadores a diferentes niveles, desde el submolecular al ecológico. Sin embargo, incluso entre aquellos investigadores que no han levantado la pancarta de la Vida Artificial se reconoce generalmente que la mayor parte de su investigación teórica sobre la evolución —por ejemplo, la de los recientes trabajos discutidos en este libro— hubiera sido sencillamente impensable sin las simulaciones por ordenador, para someter a prueba (confirmar o no confirmar) las intuiciones de los teóricos. Como hemos visto, la propia idea de la evolución como un proceso algorítmico no ha podido ser formulada y evaluada de modo apropiado hasta que ha sido posible someter a prueba enormes y complicados modelos de algoritmos, en lugar de los primitivos y excesivamente simples modelos aplicados por los primeros teóricos.

Ahora, algunos problemas científicos no son adecuados para soluciones por simulación y otros, probablemente, solamente son apropiados para soluciones por simulación pero, entre medio, hay problemas que pueden, en principio, ser tratados por ambas vías, reminiscencia de las dos diferentes vías para resolver el problema del tren planteado por Von Neumann: una vía «profunda», a través de la teoría, y una vía «superficial», mediante la fuerza bruta de la simulación y de la inspección. Sería una vergüenza que las muchas atracciones innegables de mundos simulados ahogaran nuestras aspiraciones en comprender estos fenómenos a través de la vía «profunda» de la teoría. En cierta ocasión, tuve la oportunidad de hablar con Conway sobre la creación del Juego de la Vida, y se lamentaba de que las exploraciones del mundo de la vida fueran ahora casi exclusivamente métodos «empíricos», mediante la disposición de todas las variaciones de interés sobre un ordenador y dejándolo correr para observar lo que sucede. Esta actitud no sólo evita la oportunidad de diseñar una prueba estricta de lo que uno encuentra, sino que, señalaba Conway, los que utilizan la simulación en el ordenador suelen ser muy poco pacientes; ensayan combinaciones y las observan durante quince o veinte minutos, y si durante este tiempo no sucede nada de interés, las abandonan, calificándolas como vías exploradas y estériles. Este estilo miope de exploración corre el riesgo de cerrar prematuramente importantes vías de investigación. Este es un accidente de trabajo en todas las simulaciones por ordenador, como si fuera la versión en la alta tecnología de la debilidad fundamental del filósofo: confundir un fallo de la imaginación en un proceso de discernimiento con una necesidad. Una imaginación incrementada protésicamente con un ordenador es también susceptible de fallar, sobre todo si no se utiliza con suficiente rigor.

Ahora ha llegado el momento de volver a mi cuestión principal. Cuando Conway y sus estudiantes decidieron en primer lugar crear un mundo bidímensional en el que sucedieran cosas interesantes, se encontraron que nada parecía funcionar. Este industrioso grupo de inteligentes investigadores tardó más de un año en hallar la simple ley física de la vida en el «vasto» espacio de todas las reglas posibles. No hallaron nada esperanzador en las variaciones obvias. Para conseguir darle algún sentido a este mundo artificial, trataron de alterar las constantes para el nacimiento y la muerte —cambiando, por ejemplo, la regla del nacimiento de tres a cuatro— y observaron lo que sucedía. Los mundos gobernados por estas variaciones o bien se solidificaban inmediatamente o bien se evaporaban en la nada también inmediatamente. Conway y sus estudiantes querían conseguir un mundo en el que fuera posible

el crecimiento, pero no demasiado explosivo; un mundo en el que las «cosas» —patrones de casillas de orden elevado— pudieran moverse y cambiar, pero reteniendo también su identidad a lo largo del tiempo. Y naturalmente, tendría que ser un mundo en el que las estructuras pudieran «hacer cosas» de interés (como comer, hacer recorridos o repeler cosas). De todos los imaginables mundos bidimensionales, hasta donde Conway conoce, sólo hay uno que cumple estos *desiderata*: el mundo de la vida. En cualquier caso, las variaciones del mundo de la vida que han sido comprobadas en los años siguientes nunca han estado cerca de poder medirse, en términos de interés, con los resultados de Conway, en lo que se refiere a simplicidad, fecundidad y elegancia. El mundo de la vida es el mejor de todos los mundos posibles (bidimensionales).

Supongamos ahora que algunas de las máquinas auto-reproductoras de Turing entablaran una conversación en el mundo de la vida con cada una de las otras máquinas sobre el mundo en el que se encontraban, con su admirable y simple física; expresable en una única sentencia y cubriendo todas las eventualidades[58]. Estarían cometiendo un error lógico si argumentaran que, dado que ellas [las máquinas] existen, el mundo de la vida, con su particular ley física, *tendría* que existir; después de todo, Conway podía haber decidido hacerse fontanero o jugar al *bridge* en lugar de ir de caza. ¿Qué pasaría si ellas (las máquinas autorreproductoras) dedujeran que el mundo era demasiado admirable, con su elegante física mantenedora de la vida, para que hubiera llegado a existir sin una Inteligencia Creadora? ¡Si *llegasen* a la conclusión de que debían su existencia a las actividades de un sabio Legislador, estarían en lo cierto! Hay un Dios y su nombre es Conway.

Con todo, las máquinas habrían llegado a una conclusión. La existencia de un universo que obedeciese a una serie de leyes aunque fueran tan elegantes como las leyes de la vida (o las leyes de nuestra propia física) no requiere lógicamente un Legislador inteligente. Recuérdese, en primer lugar, que la historia actual del Juego de la Vida divide la labor intelectual en dos partes: por un lado, el trabajo de exploración inicial que conduce a la ley física promulgada por el Legislador y, por otro, el trabajo de ingeniería de los explotadores de la ley, los artífices. Esto *puede* haber sucedido en este orden temporal: primeramente Conway, en un rapto de inspiración genial, promulga la física del mundo de la vida, y después él y sus estudiantes diseñan y construyen los admirables habitantes de este mundo de acuerdo con la ley establecida. Pero, de hecho, las dos partes se entremezclan: muchos ensayos de «prueba y error» para conseguir cosas que fueran interesantes facilitaron

una guía para la búsqueda de la ley apropiada. En segundo lugar, esta postulada división del trabajo ilustra un tema darwiniano fundamental del capítulo anterior. La tarea del Dios sabio necesaria para poner en movimiento a este mundo es una tarea de descubrimiento, no de creación, un trabajo para Newton, no para Shakespeare. Lo que Newton encontró —y también Conway— fueron eternos puntos fijos platónicos que cualquiera podía, en principio, haber descubierto, y no una creación idiosincrásica que depende, en cualquier caso, de las particularidades de las mentes de sus autores. Si Conway no se hubiese dedicado nunca a diseñar mundos de autómatas celulares, si Conway no hubiera existido nunca, otros matemáticos hubieran podido muy bien acertar con el mundo de la vida, del cual Conway merece el crédito. Así, cuando seguimos el camino de Darwin, Dios el Artífice se convierte en Dios el Legislador, que ahora puede ser considerado como fundiéndose con Dios *el que encuentra* la Ley. ¡La hipotética contribución de Dios llega a ser, de este modo, menos personal y, en consecuencia, más fiablemente realizable por algo persistente y no inteligente!

Hume ya había demostrado cómo funciona el argumento y ahora, apoyados por nuestra experiencia con el pensamiento darwiniano en un terreno que nos es más familiar, podemos extrapolar una alternativa darwiniana positiva a la hipótesis de que nuestras leyes son un regalo de Dios. ¿Cuál debería ser la alternativa darwiniana? Que ha habido una evolución de mundos (en el sentido de la totalidad del universo) y que el mundo en el que nos encontramos es simplemente uno entre otros incontables que han existido desde la eternidad. Existen dos vías de pensamiento, completamente diferentes sobre la evolución de las leyes, una de ellas, más fuerte, más «darwiniana» que la otra, puesto que implica la selección natural.

¿Es posible que haya habido algún tipo de *reproducción diferencial* de los universos, con algunas variantes que hayan tenido más «progenie» que otras? El Philo de Hume acarició esta idea, como ya vimos en el capítulo I:

> ¿Qué sorpresa nos invadiría si lo que encontramos es un estúpido mecánico que imita a los otros y copia un arte, que, a través de una larga sucesión de edades, después de múltiples ensayos, errores, correcciones, deliberaciones y controversias, ha sido gradualmente mejorado? Muchos mundos pueden haber sido construidos chapuceramente, a través de una eternidad, antes de que su sistema fuera destruido: mucho trabajo perdido: muchos ensayos infructuosos llevados a cabo: una lenta pero continuada mejoría conseguida durante las infinitas edades de la construcción del mundo (*Diálogos*, V).

Hume imputa la «continua mejora» a la predisposición mínimamente selectiva de un «estúpido mecánico», pero nosotros podemos sustituir al estúpido mecánico por algo aún más estúpido, sin disipar el poder elevador:

un proceso algorítmico darwiniano que va ensayando el mundo. Aunque Hume, obviamente, no pensó que esto fuera algo más que una divertida fantasía filosófica, la idea ha sido recientemente desarrollada en algún detalle por el físico Lee Smolin (1992). La idea básica es que las singularidades conocidas como *agujeros negros* son, en efecto, los lugares de nacimiento de la progenie del universo, progenie en la que las constantes fundamentales de la física difieren ligeramente, de manera aleatoria, de las constantes físicas del universo original. De este modo, de acuerdo con la hipótesis de Smolin, tendremos, al mismo tiempo, la reproducción diferencial y la mutación, las dos características esenciales de cualquier algoritmo darwiniano de selección. Aquellos universos en los que sucedió que sus constantes físicas potenciaban el desarrollo de agujeros negros habrían tenido *ipso facto* más progenie, las cuales tendrían más progenie y así sucesivamente; este es el paso de la selección. En este escenario no aparece una Parca para universos; todos los universos viven y mueren a su debido tiempo, pero algunos tienen más progenie. De acuerdo con esta idea, no es una mera e interesante coincidencia que vivamos en este universo en el que hay agujeros negros, ni es una necesidad lógica absoluta. Es, en cierto modo, el tipo de necesidad casi condicional que encontramos en cualquier explicación evolutiva. El lazo de unión, sostiene Smolin, es el carbón, que desempeña un papel tanto en el colapso de las nubes gaseosas (o, en otras palabras, el nacimiento de las estrellas, precursor del nacimiento de los agujeros negros) como en nuestra ingeniería molecular.

¿Se puede probar esta teoría? Smolin ofrece algunas predicciones que, si no se confirman, eliminarían completamente su idea: este sería el caso de la predicción de que todas las «casi» variaciones en las constantes físicas de los valores que disfrutamos deben originar universos en los que los agujeros negros son menos probables o menos frecuentes que en nuestro propio mundo. En concreto, piensa Smolin, nuestro universo debe mostrar un *optimun*, al menos local si no global en la competición por originar agujeros negros. El problema es que hay muy pocas limitaciones, hasta lo que puedo ver, sobre qué debe considerarse como una «casi» variación y por que, aunque quizás una elaboración posterior de la teoría aclararía este punto. No es necesario decir que es difícil saber todavía lo que puede hacerse con esta idea, pero cualquiera que sea el eventual veredicto de los científicos, la idea ya sirve para asegurar una cuestión filosófica. Freeman Dyson y Fred Hoyle, entre otros muchos, piensan que existe un modelo admirable en las leyes de la física; si ellos o cualquier otro cometieran el error táctico de plantear la

retórica pregunta «¿Qué otro sino Dios podría *posiblemente* explicar esto?». Smolin daría una bonita y deflacionadora respuesta. (Yo aconsejo a mis estudiantes que desarrollen hipersensibilidad frente a las preguntas retóricas en filosofía. Sólo sirven para encubrir, empapelando la pared, las grietas que se encuentran en sus argumentos).

Pero supongamos, por el puro placer de discutir, que las especulaciones de Smolin son incorrectas; supongamos que la *selección* de los universos no funciona en absoluto. Existe una especulación semidarwiniana menos convincente que también responde hábilmente a esta retórica pregunta. Hume también jugó con esta idea menos convincente, como ya hemos señalado, en sus *Diálogos*:

> En lugar de una presunta materia infinita, como suponía Epicuro, supongamos que es finita. Un número finito de partículas es sólo susceptible de transposiciones finitas. Y debe suceder que, en una duración eterna, todo posible orden o posición debe ser logrado un infinito número de veces…
>
> Supongamos… que la materia fuera lanzada a cualquier posición por una tuerza ciega y sin guía; es evidente que su primera posición debe ser, con toda probabilidad, la más confusa y más desordenada que pueda ser imaginada, sin ninguna semejanza con aquellas obras de la invención humana, en las que, además de una simetría de las partes, se descubre un ajustamiento de los medios a los fines y una tendencia a la autopreservación… Supongamos que la fuerza actuante, cualquiera que sea, aún continúa en funcionamiento… Así, el universo marcharía a través de muchas edades en una continua sucesión de caos y desorden. Pero ¿no es posible que esa situación termine por estabilizarse? ¿No deberíamos esperar esa situación, o incluso estar seguros de ella, a partir de las eternas revoluciones de la materia sin guía, y no podría esto explicar toda la aparente sabiduría e inventiva que existe en el universo? (*Diálogos*, VIII).

Esta idea no promueve ninguna versión del proceso de la selección, sino que simplemente llama nuestra atención sobre el hecho de que tenemos que jugar eternamente. En este caso, no existe una fecha límite de cinco mil millones de años, como en la evolución de la vida sobre la Tierra. Como vimos en nuestras consideraciones sobre las bibliotecas de Babel y de Mendel, necesitamos reproducción y selección si estamos dispuestos a atravesar los «vastos» espacios en una «no vasta» cantidad de tiempo, pero cuando el tiempo deja de ser una consideración limitadora, la selección ya no es una exigencia. En el curso de la eternidad, podemos ir *a cualquier parte* en la Biblioteca de Babel o de Mendel —o en la Biblioteca de Einstein [que incluiría todos los posibles valores de todas las constantes de la física]— siempre que nos mantengamos en movimiento. (Hume imagina una «fuerza actuante» para mantener la dificultosa marcha y esto nos recuerda el argumento de Locke sobre la materia sin movimiento, aunque no supone que la «fuerza actuante» tenga inteligencia alguna). En efecto, si nos movemos con dificultad a través de todas las posibilidades durante la eternidad,

¡pasaremos a través de todos los posibles lugares en estos «vastos» (pero finitos) espacios, no una vez sino una infinidad de veces!

En años recientes, físicos y cosmólogos han considerado en serio varias versiones de esta especulación. John Archibald Wheeler (1974), por ejemplo, ha propuesto que el universo se balancea hacia atrás y hacia delante a lo largo de la eternidad: al Big Bang le sigue una expansión, a la cual le sigue una contracción o aplastamiento, tras el cual sigue otro Big Bang y así sucesivamente en cada oscilación con variaciones aleatorias en las constantes físicas y en otros parámetros cruciales. Cada posible escenario es ensayado una infinidad de veces y, de este modo, cada variación sobre cada tema, tanto las que «tienen sentido» como las que son absurdas, tiene lugar no una, sino una infinidad de veces.

Resulta difícil creer que esta idea pueda ser comprobada empíricamente de un modo inteligible, aunque debamos reservarnos la opinión. Variaciones o elaboraciones sobre el tema pueden tener implicaciones que cabe confirmar o rechazar. Mientras tanto, vale la pena subrayar que esta familia de hipótesis tiene la virtud de extender los principios de la explicación que funciona tan bien en los dominios de las verificaciones, hasta sus últimas consecuencias. Consistencia y simplicidad están a su favor. Y todo esto, de nuevo, es suficiente para frenar la atracción de la alternativa tradicional[59].

Quienquiera que hubiese ganado un torneo de lanzamiento de monedas a cara o cruz podría haber caído en la tentación de creerse poseído por poderes mágicos, especialmente si no conociera directamente a los otros jugadores. Supongamos que decidimos organizar un torneo de diez rondas de lanzamiento de monedas a cara o cruz, sin permitir que ninguno de los 1.024 «concursantes» se den cuenta de que van a participar en un torneo. A cada uno de los jugadores les decimos en el momento de reclutarlo: ¡felicidades, amigo mío! Yo soy Mefistófeles y voy a otorgarte grandes poderes. ¡A mi lado vas a ganar diez veces consecutivas el lanzamiento de monedas sin perder una sola vez! Nos las arreglamos para que nuestros concursantes se reúnan de dos en dos, hasta que al final tengamos un ganador. (¡Nunca permitiremos que los jugadores discutan su relación entre sí y despediremos a los 1.023 perdedores con alguna burla *sotto voce* para crear el efecto de que eran tan estúpidos como para creerse la afirmación de que éramos Mefistófeles!). Al ganador —sólo uno— lo someteremos a alguna prueba de ser un escogido, y si se lo cree, será en realidad una ilusión de lo que podría llamarse una miopía retrospectiva. El ganador no se ha dado cuenta de que la

situación estaba estructurada de tal modo que alguien tenía que ser el afortunado, y justamente le tocó a él.

En este sentido, si el universo estuviera estructurado de tal modo que una infinidad de diferentes «leyes físicas» fueran comprobadas a lo largo de los tiempos, sucumbiríamos a la misma tentación si extrajésemos la conclusión de que las leyes de la naturaleza han sido preparadas especialmente para nosotros. Este no es un argumento para llegar a la conclusión de que el universo es, o debe ser, estructurado de tal manera, sino a la conclusión más modesta de que ninguna de las características de las observables «leyes de la naturaleza» puede ser invulnerable a esta alternativa interpretación deflacionista.

Una vez formuladas estas hipótesis darwinianas, aún más especulativas e incluso más atenuadas, sirven —a la clásica manera darwiniana— para facilitar, en pequeños pasos, la tarea explicativa a la que nos enfrentamos. Todo lo que queda por explicar en este punto es una perceptible elegancia o prodigiosidad en las leyes de la física observadas. Si tenemos dudas acerca de que la hipótesis de un infinito número de variantes del universo pueda realmente explicar esta elegancia, deberíamos revelar que ésta tiene al menos la misma exigencia de ser una cuestión que no necesita explicación, que la alternativa tradicional; en un tiempo en el que Dios ha sido despersonalizado hasta el punto de ser un tipo de principio abstracto de belleza y bondad, es difícil admitir que la existencia de Dios pueda explicar algo. ¿Qué se afirmaría mediante la «explicación» que no haya sido aportado en la descripción de los maravillosos fenómenos que han de ser explicados?

Darwin inició sus ataques en la parte media de la pirámide cósmica: dadme orden y tiempo y explicaré el diseño. Hemos visto ahora cómo el ácido universal fluye cuesta abajo: si otorgamos a los sucesores de Darwin el caos (en el viejo sentido, pasado de moda, de pura aleatoriedad sin sentido) y la eternidad, ¿podrán explicar el orden, el mismo orden necesario para explicar el diseño? A su vez, ¿*necesita* el caos absoluto una explicación? ¿Qué queda por explicar? Algunos piensan que sigue sin plantearse el «por qué» de otra cuestión: *¿Por qué hay algo y no nada?* Las opiniones difieren sobre si la pregunta es, en último término, una pregunta inteligible[60]. Si lo fuera, la respuesta «porque Dios existe» es probablemente tan buena como cualquier otra, pero prestemos atención a la respuesta competidora: ¿Por qué no?

4. *El eterno retorno. ¿Vida sin fundamentos?*

> La ciencia es admirable cuando destruye respuestas metafísicas, aunque sea incapaz de aportar otras para sustituirlas. La ciencia elimina los fundamentos sin reemplazarlos. Tanto si deseamos estar allí como si no, la ciencia nos coloca en la posición de tener que vivir sin fundamentos. Fue horrible cuando Nietzsche lo dijo aunque actualmente es lo habitual; nuestra posición histórica —y no hay fin a la vista — es la de tener que filosofar sin «fundamentos».
>
> Hilary Putnam, *Las mil caras del realismo*

> La sensación de que el significado del universo se había evaporado era lo que parecía escapársele a aquellos que dieron la bienvenida a Darwin como un benefactor de la humanidad. Nietzsche consideraba que la evolución presentaba una visión correcta del mundo, pero que era una visión desastrosa. Su filosofía era un intento de conseguir una nueva visión del mundo que tuviera en cuenta el darwinismo pero que no fuese destrozada por éste.
>
> R. J. Hollingdale, *Nietzsche: The Man and his Philosophy*

Como resultado de la publicación de *El origen de las especies*, Friedrich Nietzsche redescubrió lo que Hume ya había acariciado: la idea de que una eterna repetición de una variación ciega e insensata —torpes movimientos caóticos y sin sentido de materia y de ley— originaría inevitablemente mundos cuya evolución a través del tiempo generaría las *aparentemente* absurdas historias de nuestras vidas. Esta idea del eterno retorno ha llegado a ser la piedra angular de su nihilismo y, de esta manera, el fundamento en parte de lo que ha llegado a ser el existencialismo.

La idea de que lo que está sucediendo ahora ha estado sucediendo antes debe ser tan vieja como el fenómeno del *déjà-vu*, que tan a menudo inspira versiones supersticiosas. Las cosmogonías cíclicas no son inusuales en el catálogo de las culturas humanas. Pero cuando a Nietzsche se le ocurre una versión de la concepción de Hume —y de John Archibald Wheeler— la adopta como mucho más que un divertido experimento mental o una elaboración de antiguas supersticiones. Nietzsche piensa —al menos durante un tiempo— que había encontrado accidentalmente una prueba científica de la mayor importancia[61]. Sospecho que Nietzsche fue estimulado a tomar la idea más seriamente de lo que Hume había hecho debido a su insuficiente apreciación del tremendo poder del pensamiento darwiniano.

Las referencias de Nietzsche a Darwin son casi siempre hostiles, aunque pocas, y esto apoya el argumento de Walter Kaufmann en el sentido de que «Nietzsche no era un darwinista aunque sólo fuera despertado de su dogmático sopor por Darwin, de modo parecido a como Kant lo fue, un siglo antes, por Hume» (Kaufmann 1950). Las referencias de Nietzsche a Darwin

también revelan que sus relaciones con las ideas de éste estuvieron dificultadas con frecuentes tergiversaciones y malentendidos, ya que quizá «conoció» primero a Darwin a través de la apropiación entusiasta de sus muchos divulgadores en Alemania y, naturalmente, a través de Europa. En los escasos puntos de crítica específica en los que Nietzsche se aventura, éste desconoce la posibilidad de una «selección inconsciente», cuando ésta es una de las más importantes ideas fundamentales en el *El origen de las especies*. Nietzsche se refiere a la «completa *bêtise* de los ingleses Darwin y Wallace» y se queja de que «al final, la confusión llega tan lejos que uno considera al darwinismo como filosofía: ahora los sabios y los científicos dominan» (*La voluntad de poder*, 1901). Otros, sin embargo, consideran que Nietzsche era un darwinista —«Otros sabios castrados han sospechado que yo era un darwinista por causa de él» (*Ecce homo*, 1889)—, una etiqueta de la que él se burlaba cuando procedía a escribir, en *La genealogía de la moral* (1887), una de las primeras y aun más sutiles de las investigaciones darwinianas sobre la evolución de la ética, un tema al que volveremos en el capítulo 16.

Nietzsche creía que su argumento sobre el eterno retorno era una prueba de lo absurdo e ininteligible de la vida, una prueba de que desde lo alto no le fue *dado* un significado al universo. Y esta es indudablemente la raíz del temor que muchos experimentan cuando se encuentran con Darwin, de modo que vamos a examinar esta cuestión en una versión de Nietzsche, tan extrema como hemos sido capaces de encontrar. ¿Por qué, exactamente, el eterno retorno haría la vida ininteligible? ¿No es esto obvio?

Reflexionemos sobre lo que pasaría si un demonio se arrastrara detrás de uno de nosotros un día o una noche, y nos dijera, en nuestra soledad: «Esta vida que vives y has vivido, debes vivirla otra vez e innumerables veces más. Y no habrá nada nuevo en ella, ya que cada dolor y cada alegría y cada pensamiento y cada mirada —todo lo inexpresablemente pequeño y grande en vuestra vida— debe repetirse de nuevo en ti y con la misma frecuencia y series…». ¿No nos lanzaríamos al suelo gritando maldiciones al demonio que así nos habló? O, por el contrario, disfrutaríamos de un momento tremendo en el que le responderíamos: «¡Tú eres un dios y nunca he escuchado a alguien más divino!» (*La gaya ciencia*, 1882.).

¿Es este mensaje liberador u horripilante? Nietzsche no parecía aceptarlo, quizá debido a que, a menudo, escogía para vestir las implicaciones de sus «más científicas hipótesis» esos adornos bastantes místicos. Podemos introducir un poco de aire fresco en la discusión si consideramos una deliciosa parodia del novelista Tom Robbins en *También las vaqueras sienten melancolía*, 1976:

En la Navidad de aquel año, Julian regaló a Sissi una miniatura que representaba un pequeño pueblo tirolés. El trabajo de artesanía era extraordinario.

Había una pequeña catedral cuyas vidrieras de colores se convertían en ensaladas de frutas a la luz del sol. Había una plaza y un *Biergarten*, El *Biergarten* era muy ruidoso los sábados por la noche. Había un ayuntamiento y una comisaría de policía, con secciones que la hacían visible desde el exterior y que mostraban las cantidades habituales de papeleo burocrático y de corrupción. Había pequeños tiroleses con pantalones de cuero, anudados de manera intrincada y debajo de los pantalones genitales tallados con la misma habilidosa técnica. Había tiendas de esquí y muchas otras cosas interesantes, incluyendo un orfanato. El orfanato estaba diseñado para prenderle fuego y quemarlo todos los días de Navidad. Los huérfanos se lanzarían a la nieve con sus pijamas envueltos en llamas. Terrible. Alrededor de la segunda semana de enero, un bombero vendría y buscaría entre las ruinas murmurando: «Si simplemente me hubieran escuchado, estos niños estarían hoy vivos».

La destreza con la que está escrito este pasaje es realmente extraordinaria. La repetición del drama del orfanato todos los años parece privar al pequeño mundo de todo significado. Pero ¿por qué? ¿Por qué exactamente debe ser el lamento del bombero por la repetición del fuego lo que hace que todo suene tan vacío? Quizá si miramos más de cerca lo que esto acarrea encontraremos el truco que hace que el relato «funcione». ¿Son los pequeños tiroleses los que tienen que reconstruir el orfanato o existe un botón que dice RESET (repetir) en este pueblo de miniatura? ¿Qué diferencia existe entre una u otra opción? ¿De dónde vendrán los nuevos huérfanos? ¿Volverán los huérfanos «muertos» a la vida? (Dennett 1984:9-10). Obsérvese que Robbins dice que el orfanato estaba diseñado para empezar a arder y quemarse cada día de Navidad. El creador de este mundo en miniatura se está, evidentemente, burlando de nosotros, ridiculizando la seriedad con la que nos enfrentamos a estos problemas de la vida. La moraleja parece clara: si el significado de este drama viene desde lo alto, de un Creador, se trataría de un chiste obsceno, una trivialización de los esfuerzos de los individuos en este mundo. Pero ¿qué pasaría si el significado es, de algún modo, la creación de los individuos por sí mismos, como surgiendo de nuevo en cada encarnación, más que un regalo venido del cielo? Esta alternativa puede abrir la posibilidad de un significado para aquel pequeño mundo que no fuera amenazado por la repetición del fuego cada Navidad.

Este es el tema que define al existencialismo en sus diferentes especies: el único significado que podría haber es el significado de que nosotros nos creamos a nosotros mismos. Cómo puede ser realizado este *truco* ha sido siempre casi un misterio entre los existencialistas pero, como veremos pronto, el darwinismo tiene alguna desmitificación que ofrecer en su visión del proceso de una creación con sentido. La clave, de nuevo, es el abandono de la concepción de la primera mente de Locke y su sustitución por una concepción en la que la *importancia en sí misma*, como todo aquello que atesoramos, evoluciona gradualmente desde la nada.

Podemos hacer una pausa, antes de volver a algunos de estos detalles, para reconsiderar a dónde nos ha llevado nuestra tortuosa ruta. Hemos empezado con una visión infantil de un antropomórfico Dios Artesano y hemos reconocido que esta idea, tomada literalmente, se encontraba en vías de extinción. Cuando miramos a través de los ojos de Darwin al proceso real de diseño del que nosotros y todas las maravillas de la naturaleza somos, hasta la fecha, sus productos, nos damos cuenta de que Paley estaba en lo cierto cuando veía estos efectos como el resultado de una gran cantidad de trabajo de diseño, aunque encontramos una explicación nada milagrosa: un proceso algorítmico de diseño, no inteligente, masivamente paralelo y, en consecuencia, prodigiosamente derrochador, en el que, sin embargo, los mínimos incrementos de diseño han sido utilizados económicamente, copiados y vueltos a copiar a lo largo de miles de millones de años. La admirable *particularidad* o *individualidad* de la creación se debía, no a la inventiva genial de Shakespeare, sino a la incesante contribución del azar, una secuencia de crecimiento a la que Crick [1968] denominó «accidentes congelados».

Esta concepción del proceso creativo deja aún en apariencia un papel para Dios como legislador, pero esta alternativa concedida con el newtoniano papel del legislador también se ha evaporado, como hemos visto recientemente, no dejando detrás ningún agente inteligente que participe en el proceso. Lo que queda es que este proceso, moviéndose con torpeza a través de la eternidad, encuentra, sin inteligencia (cuando encuentra algo) una eterna posibilidad platónica de orden. Esto es, naturalmente, algo bello, como los matemáticos exclaman siempre, pero no es en sí mismo algo inteligente sino, maravilla de las maravillas, algo inteligible. Ser abstracto y fuera del tiempo es la nada con una *iniciación* u *origen* que necesita explicación**[62]**. Lo que necesita que se explique su origen, es el origen del universo concreto, como el Philo de Hume hace mucho tiempo preguntaba: ¿por qué no detenerse en el

mundo material? Esto, como hemos visto, es una versión del engaño definitivo que dice que no necesita nada de nadie; se crea a sí mismo *ex nihilo* o se crea a cualquier ritmo a partir de algo que es casi indistinguible de la nada. Al contrario de la misteriosa creación eterna, como un rompecabezas de Dios, esta autocreación es una destreza nada milagrosa que nos ha dejado muchas huellas. Y al no ser precisamente una creación concreta, sino el producto de un proceso histórico muy particular, es una creación absolutamente única —que incluye y empequeñece como creaciones todas las novelas, cuadros y sinfonías de todos los artistas— que ocupa un lugar en el hiperespacio de posibilidades que difieren unas de otras y de todas las demás.

Benedicto Spinoza, en el siglo XVII, *identificó* a Dios con la naturaleza argumentando que la investigación científica era el verdadero camino de la teología: por esta herejía fue perseguido. En la herética visión de Spinoza, *Deus sive Natura* (Dios o Naturaleza), hay una turbadora cualidad, algo así como las dos caras de Jano: al proponer su simplificación científica, ¿estaba Spinoza personificando a la naturaleza o despersonalizando a Dios? La concepción de Darwin, más generativa, aporta la estructura en la que podemos ver la inteligencia de la madre naturaleza (¿o es meramente inteligente?) como una característica de algo que se crea a sí mismo, nada milagrosa y nada misteriosa y, en consecuencia, todavía más admirable.

CAPÍTULO 8
La biología es ingeniería

I. *Las ciencias de lo artificial*

> Desde la Segunda Guerra Mundial los descubrimientos que han cambiado el mundo no se han realizado tanto en las soberbias instalaciones de la física teórica como en los menos conocidos laboratorios de la ingeniería física y experimental. Los respectivos roles de la ciencia pura y de la ciencia experimental han cambiado; la edad de oro de la física ya no es lo que era, la edad de Einstein, Schrödinger, Fermi y Dirac… Los historiadores de la ciencia han tenido a bien ignorar la historia de los grandes descubrimientos en la física aplicada, ingeniería y ciencia de la computación, donde el progreso científico real debe ser encontrado en estos tiempos. La ciencia de la computación, en especial, ha cambiado y continúa cambiando la cara del mundo más completa y más drásticamente que cualquiera de los grandes descubrimientos de la física teórica.
>
> Nicholas Metropolis,
> *The Age of Computing: A Personal Memoir*

En este capítulo deseo seguir la pista de las implicaciones que ha tenido un hecho central —me aventuro a decir *el* hecho central— de la revolución darwiniana, y que han pasado inadvertidas o menospreciadas: el matrimonio, según Darwin, entre la biología y la ingeniería. Mi objetivo en este capítulo es describir el lado positivo de la historia de la biología como ingeniería. En capítulos siguientes nos ocuparemos de varios asaltos y retos, aunque antes de que éstos asuman el protagonismo, quiero exponer el argumento de que la biología, vista desde la perspectiva de la ingeniería, no es simplemente útil en ocasiones, no es una simple opción valiosa, sino el organizador obligatorio de todo el pensamiento darwiniano y la fuente primaria de su poder. Ante esta afirmación, espero una razonable resistencia emocional. Pero que sea honesta: que el título de este capítulo no provoque una reacción negativa en mis lectores, expresada más o menos con el siguiente tono: «¡Oh no! ¡Qué comparación tan deprimente, filistea, y reduccionista! ¿Es *bastante* más importante la biología que la ingeniería?».

La idea de que un estudio de los seres vivos es, al menos, algo cercano a la ingeniería ha sido moneda corriente desde las pioneras investigaciones de

Aristóteles sobre los organismos y sus análisis de la teleología, la cuarta de sus causas, aunque únicamente a partir de Darwin se le ha prestado atención. La idea está completamente explícita en el argumento del diseño, el cual invita al observador del mundo a maravillarse ante la ingeniosa interrelación entre las partes, la planificación elegante y la técnica artesana exquisita del Artífice. Pero la ingeniería ha tenido siempre un estatus de segunda clase en el mundo intelectual. Desde Leonardo da Vinci a Charles Babbage y a Thomas Edison, en la ingeniería el genio ha sido siempre aclamado aunque, con todo, mirado con cierto grado de condescendencia por la élite de los mandarines de la ciencias y las artes. Aristóteles no ayudó a mejorar la situación al proponer una distinción, adoptada por la filosofía medieval, entre lo que era *secundum naturam* —de acuerdo con la naturaleza— y lo que era *contra naturam* —contra la naturaleza o artificial. Los mecanismos, pero no los organismos, eran *contra naturam*. También existían aquellas cosas que eran *praeter naturam*, o *in*naturales (monstruos y mutantes), y las cosas que eran *super naturam* como los milagros (Gabbey 1993). ¿Cómo puede el estudio de lo que era *contra* la naturaleza iluminar las glorias, incluso los monstruos y los milagros, de la naturaleza?

Las huellas fósiles de esta actitud negativa se encuentran por doquier en nuestra cultura. Por ejemplo, en mi propia disciplina, la filosofía, la subdisciplina conocida como filosofía de la ciencia tiene una historia larga y respetable; muchos de los filósofos más eminentes e influyentes de la actualidad son filósofos de la ciencia. Hay excelentes filósofos de la física, de la biología, de las matemáticas e incluso de las ciencias sociales. Pero nunca he oído hablar de nadie en el campo de la filosofía que se dedicara a la filosofía de la ingeniería, como si no hubiera bastante materia en este campo, de interés conceptual, para que un filósofo se especializara en ella. Aunque esta actitud está cambiando, cada día más y más filósofos reconocen que la ingeniería acoge en su seno algunos de los pensamientos más profundos, más bellos y más importantes nunca desarrollados. (El título de esta sección está tomado del libro de Herbert Simon (1969) dedicado a estos temas).

La gran visión de Darwin fue que todos los diseños de la biosfera podrían ser los productos de un proceso que era tan paciente como insensato, una «automática» y gradual «elevación» en el espacio de diseño. De manera retrospectiva, podemos ver que el propio Darwin difícilmente podía haber imaginado, basado tan sólo en la evidencia de que disponía, los refinamientos y las extensiones de su idea que había permitido más tarde a los darwinistas ir más allá de su propio y cauteloso agnosticismo sobre los orígenes de la misma

vida, y aun del «diseño» del orden físico que su idea presuponía. No se encontraba Darwin en mejor situación para caracterizar en el orden que iba a describir las limitaciones y los poderes del mecanismo hereditario; precisamente se dio cuenta de que tenía que existir tal mecanismo y que éste había de promocionar el orden, cualquiera que fuese, que hiciera que el «descenso con modificación» no sólo fuera posible sino fructífero.

El transcurso de más de un siglo de subsiguientes enfoques y extensiones de la gran idea de Darwin ha ido acompañado puntualmente por la controversia, ampliamente ilustrada, de paso, por la reflexiva extensión de la idea de Darwin sobre sí misma: la evolución de los *memes* darwinianos acerca de la evolución no ha sido solamente acompañada sino positivamente acelerada por la competición entre ideas. Y de acuerdo con la misma hipótesis que Darwin planteó con respecto a los organismos, «la competición será por lo general más fuerte entre aquellas formas que se encuentran más relacionadas entre sí» (*El origen de las especies*). Los propios biólogos no han resultado inmunes, naturalmente, a la herencia de estas actitudes negativas sobre la ingeniería. Después de todo, ¿cuál es el objetivo anhelado respecto a lo que hemos llamado «ganchos celestes», sino la ingenua esperanza de que, de algún modo, se produzca un milagro que eleve el proceso evolutivo por encima de las denominadas «grúas»? Una persistente resistencia subliminal a este aspecto de la idea fundamental de Darwin ha aumentado la controversia, impidiendo su comprensión y distorsionando su expresión, mientras que, al mismo tiempo, lanzaba al darwinismo algunos de los retos más importantes.

Como respuesta a estos retos, la idea de Darwin ha seguido creciendo y haciéndose más fuerte. Hoy podemos ver que no sólo las distinciones de Aristóteles sino también otras estimadas compartimentaciones de la ciencia están amenazadas por su expansión territorial. Los alemanes dividen el saber en *Naturwissenschaften* (‘ciencias de la naturaleza’) y *Geisteswissenschaften* (‘ciencias del espíritu’), pero esta tajante división —casi prima hermana de la utilizada en el libro de C. P. Snow (1963) titulado *Dos culturas*— se encuentra amenazada por la perspectiva de que la ingeniería se extienda desde la biología a las ciencias humanas y a las artes. Si hay un espacio de diseño en el que las progenies de nuestros cuerpos y de nuestras mentes se encuentren unidas bajo un holgado conjunto de procesos de Investigación y Desarrollo, I+D, entonces estas tradicionales barreras pueden derrumbarse.

Antes de seguir adelante, deseo enfrentarme a una sospecha. Dado que hemos admitido como probado que el propio Darwin no apreciaba muchas de

las cuestiones que deben tratarse si la teoría de la evolución por selección natural ha de sobrevivir, ¿no resulta algo trivial o tautológica mi afirmación de que la idea de Darwin sobrevive a todos estos retos? ¡No es de extrañar que pueda mantenerse o extenderse, ya que sigue cambiando como respuesta a nuevos retos! Si mi propósito fuera coronar a Darwin como autor o como héroe, merecería esta sospecha, aunque naturalmente este no es, en principio, un ejercicio de historia intelectual. Para mi tesis no importa realmente que el propio Darwin no existiese. Darwin podría ser, en lo que a mí respecta, «el contribuyente medio», una especie de mítico autor virtual. (Algunas personalidades colocan a Homero en esta categoría). El hombre histórico real me fascina; su curiosidad, integridad y perseverancia me inspiran; sus miedos personales y sus fallos lo hacen ser querido. Pero todo estas cosas respecto a Darwin, hasta cierto punto, no vienen al caso. Darwin tuvo la buena fortuna de ser el partero de una idea que tiene su propia vida, debido precisamente a que crece y cambia. Esto no sucede en la mayoría de las ideas.

De hecho, entre los respectivos partidarios de ambos frentes de la controversia se ha extendido una gran cantidad de retórica, acerca de si el propio Darwin —lo podríamos llamar Saint Charles— era un gradualista, un adaptacionista, un catastrofista, un capitalista o un feminista. Las respuestas a estas cuestiones son, por sí mismas, de considerable interés histórico, y si las separamos cuidadosamente de las cuestiones de justificación fundamental, pueden ayudarnos hoy a comprender cuáles son realmente los temas científicos. Aquello que algunos pensadores *piensan* que están haciendo —salvar al mundo de un *ismo* o de otro, encontrar un espacio para Dios en la ciencia, o combatir la superstición— a menudo está en contradicción con lo que sus campañas realmente hacen. Ya hemos visto ejemplos de este tipo, además de los que llegarán. Ningún área de la investigación científica es impulsada por más asuntos ocultos que la teoría evolucionista, y esto, en realidad, ayudará a descubrirlos, pero nada se deriva directamente del hecho que alguna gente está tratando desesperadamente, de manera consciente o inconsciente, de proteger o de destruir algún mal. A veces la gente lo consigue correctamente, a pesar de haber sido impulsada su acción por los deseos más impresentables. Darwin fue quien fue y pensó lo que pensó, verrugas incluidas. Y ahora está muerto. El darwinismo, por otro lado, tiene más de nueve vidas y se esfuerza honestamente en ser inmortal.

2. *Darwin ha muerto. ¡Viva Darwin!*

He sacado el título de esta sección del título del «Résumé» con el que Manfred Eigen termina su libro publicado en 1992. Hay un inconfundible aroma de ingeniería en el pensamiento de Eigen. Su investigación es una secuencia de problemas de construcción biológica planteados y resueltos: ¿cómo hacer que los materiales se acumulen en el lugar de la construcción y cómo hacer que el diseño quede determinado, y en qué orden deben están unidas las distintas partes de modo que no se derrumben antes de que la estructura global se complete? Su tesis es que las ideas que presenta son revolucionarias pero que, después de esta revolución, el darwinismo no sólo está vivo y en buen estado sino que sale reforzado. Me gustaría explorar este tema con más detalle, dado que veremos otras versiones que no están, de ningún modo, tan bien definidas como la de Eigen.

¿Qué es lo que se supone que es revolucionario en el trabajo de Eigen? En el capítulo 3 examinamos lo que se llama un «paisaje adaptativo» con un solo pico, y vimos como el *efecto Baldwin* podía convertir a un casi invisible poste telefónico en una llanura en el monte Fuji, con unos alrededores que se elevan progresivamente en una pendiente, así que no importa en qué espacio comenzamos, ya que alcanzaremos la cumbre si seguimos simplemente la regla local: «Nunca bajar, subir cada vez que sea posible».

La idea de un «paisaje adaptativo» fue introducida por Sewall Wright [1932] y se ha convertido en una especie de prótesis imaginativa estándar para los teóricos de la evolución. Su valor se ha probado literalmente en miles de aplicaciones, incluyendo muchas no relacionadas con la teoría de la evolución. En la inteligencia artificial, en la economía y en otros campos científicos, este modelo de resolución de problemas se ha hecho popular merecidamente con el llamado «gradiente de ascenso». Lo bastante popular como para estimular a los teóricos a calcular sus limitaciones, que son importantes. Para ciertas clases de problemas —o en otras palabras, en ciertos tipos de «paisajes»— el método del «gradiente de ascenso» es del todo impotente por una razón intuitivamente obvia: los escaladores quedan atrapados en cimas de segunda categoría en lugar de seguir buscando su camino hasta la cima global, el monte Everest de la perfección. (Las mismas limitaciones dificultan el método de la fundición simulada). La regla local es fundamental para el darwinismo; es equivalente a los requerimientos de que no puede haber ninguna previsión inteligente en el proceso de diseño, sino únicamente una explotación estúpida y oportunista de cualquier afortunada «elevación» que ocurra a lo largo del camino evolutivo.

Lo que Eigen ha demostrado es que el más simple modelo darwiniano de mejora progresiva de una única pendiente de aptitudes hacia el pico óptimo de la perfección no funciona si se trata de describir lo que sucede en la evolución molecular o viral. El ritmo de adaptación de los virus (y también de las bacterias y de otros patógenos) es mucho más rápido de lo que los modelos «clásicos» predicen, tan rápido que parece incluir ilícitos «mire hacia delante» por parte de los escaladores. ¿Significa esto que el darwinismo debe ser abandonado? De ninguna manera, porque lo que se entiende como local depende de la escala que usemos, cosa que no es sorprendente.

Eigen llamó la atención sobre el hecho de que cuando los virus evolucionan, no lo hacen en fila india; viajan en enormes multitudes de variantes casi idénticas, una nube de bordes difuminados en la Biblioteca de Mendel que Eigen denominó «casi especies». En la Biblioteca de Babel ya observamos la inimaginable enorme nube formada por las variantes de *Moby Dick*, pero cualquier biblioteca real es probable que tenga más de una o dos ediciones de un libro en sus estanterías, y en el caso de un libro realmente popular como *Moby Dick*, también es posible que existan múltiples copias de la misma edición. Como las colecciones reales de *Moby Dick*, las nubes reales de virus incluyen múltiples copias idénticas pero también múltiples copias de variaciones tipográficas menores y este hecho tiene, de acuerdo con Eigen, algunas implicaciones que han sido pasadas por alto por los darwinistas «clásicos». Es la *forma* de la nube de variantes la que tiene la clave de la velocidad de la evolución molecular.

Un término clásico entre los genetistas para denominar la versión canónica de una especie (análogo al texto canónico de *Moby Dick*) es el *tipo salvaje*. Los biólogos suponían que entre los diferentes genotipos dentro de una población, predomina el *tipo salvaje* puro. Análoga sería la afirmación de que, en cualquier biblioteca con una colección de copias de *Moby Dick*, la mayoría de las copias serían de la edición canónica (si hubiera alguna). Pero este no es el caso para los organismos, como no lo es para los libros en la bibliotecas. En efecto, el *tipo salvaje* es realmente una abstracción, como la del contribuyente medio, y una población puede no contener ningún individuo que tenga exactamente «el» *tipo salvaje* del genoma. (Esto mismo ocurre con los libros: los eruditos pueden debatir durante años acerca de la pureza de una palabra concreta en un texto concreto, y hasta que tales debates sean resueltos, nadie podrá decir exactamente cuál será el tipo canónico o el tipo salvaje de esta obra, aunque la identidad de la obra difícilmente esté en peligro. El *Ulises* de James Joyce puede ser un buen ejemplo).

Eigen puntualiza que esta distribución de la «esencia» en una variedad de vehículos idénticos hace que esta esencia sea más movible, más adaptable en especial a los abruptos «paisajes adaptativos», con múltiples picos y pocas laderas. Esto permite a la «esencia» enviar grupos de reconocimiento eficientes por las colinas y por los cerros vecinos, evitando costosas exploraciones por los valles y por lo tanto incrementando enormemente su capacidad para encontrar picos más altos, los óptimos, a alguna distancia de su centro, donde se encuentra el (virtual) tipo salvaje[63]. Las razones de este funcionamiento han sido resumidas por Eigen como sigue:

Mutantes funcionalmente competentes, cuyos valores de selección son cercanos a los del *tipo salvaje* (aunque permanecen bajo los de este tipo), alcanzan números más elevados de poblaciones que aquellos mutantes que son funcionalmente inefectivos. Se construye un espectro asimétrico de mutantes, en el que mutantes trasladados lejos del *tipo salvaje* surgen como intermediarios. En esta cadena de mutantes, la población está influida decisivamente por la estructura del «paisaje de valores». El paisaje de valores consiste en llanuras, colinas y montañas conectadas entre sí. A la altura de las montañas, el espectro de mutantes se encuentra muy esparcido y a lo largo de los cerros aparecen parientes distantes del *tipo salvaje* con frecuencia finita (es decir, no infinitesimal). Es precisamente en las regiones montañosas en las que cabe esperar la presencia de mutantes selectivamente superiores. Tan pronto como uno de estos aparece en la periferia de un espectro de mutantes el conjunto establecido se colapsa. Un nuevo conjunto se construye alrededor del mutante superior, el cual asume el papel de *tipo salvaje*… Esta cadena causal resulta de una especie de «acción en masa» mediante la que los mutantes superiores son sometidos a prueba con probabilidades más elevadas que los mutantes inferiores, aunque el segundo se encuentre a la misma distancia del tipo salvaje (Eigen 1992:25).

Hay una estrecha interacción entre la forma del *paisaje adaptativo* y la población que lo ocupa, que crea una serie de asas de *feedback* que pasan —habitualmente— de un escenario de problemas temporalmente estable a otro. Tan pronto se culmina la ascensión de un pico, el paisaje global se ondula y se levanta para formar una nueva montaña con diferente altura, y de nuevo hay que iniciar la ascensión. Si fuéramos una *casi especie* de virus, el paisaje estaría constantemente remodelándose bajo nuestros pies.

En realidad, todo esto no es tan revolucionario como Eigen afirma. El propio Sewall Wright, en su *teoría del equilibrio por remodelación*, trató de explicar cómo múltiples picos y paisajes en remodelación serían atravesables no por ejemplares de individuos de *tipo salvaje*, sino por poblaciones de tamaños diferentes compuestas por variantes; Ernst Mayr ha insistido durante años en que la *idea de población* está en el corazón del darwinismo, algo que los genetistas pasaron peligrosamente por alto. En consecuencia, Eigen no ha revolucionado en realidad el darwinismo, sino que ha creado algunas innovaciones teóricas que aclaran y refuerzan las ideas menospreciadas e imperfectamente formuladas que habían circulado durante años, lo que no es

ni mucho menos una pequeña contribución. Cuando Eigen (1992:125) dice: «¡La aceleración (cuantitativa) de la evolución que esto acarrea es tan grande que aparece al biólogo como una sorprendente nueva *calidad,* una aparente habilidad de selección para *ver por adelantado*, algo que sería considerado por los darwinistas clásicos como la más pura de las herejías!», se deja llevar por una bien conocida sobreactuación dramática, desdeñando a los muchos biólogos que, al menos, anticiparon e incluso fomentaron su «revolución».

Después de todo, cuando los teóricos darwinistas tradicionales postularon *paisajes adaptativos* y esparcieron genotipos aleatoriamente sobre ellos para calcular lo que la teoría decía que sucedería con los genotipos, averiguaron que en la naturaleza los genotipos no son lanzados aleatoriamente en partes preexistentes del mundo. Cada modelo de un proceso que consume tiempo tiene que comenzar en algún «momento» arbitrario; la cortina se levanta y el modelo entonces lleva a cabo seguidamente el plan de acción. Si observamos este modelo y vemos que al «comienzo» aparece un grupo de candidatos allá abajo, en los valles, podemos estar bastante seguros de que el teórico reconoce que estos candidatos no estarán «siempre» allí, sea lo que sea lo que esto signifique. En cualquier parte del *paisaje adaptativo* hay candidatos en un momento dado, antes había picos, si no aquellos candidatos no estarían allí, de modo que los valles que estos candidatos están ocupando deben ser relativamente nuevos, una nueva situación comprometida en la que la evolución los ha colocado. Sólo esta hipótesis puede justificar la localización de los candidatos en el primer lugar en los valles del *paisaje adaptativo*. La contribución de Eigen refuerza la apreciación de que debemos añadir estas complicaciones a los modelos si queremos que realmente hagan el trabajo que los darwinistas siempre han supuesto que sus modelos más simples pueden hacer.

Seguramente no se trata de un accidente que nuestra apreciación de la necesidad de estos modelos más complicados coincida en el tiempo (casi menos de un mes y seguro que menos de un año) con nuestra capacidad para construir y explorar tales modelos en los ordenadores existentes. No antes de que los más potentes ordenadores estuvieran disponibles, pudimos descubrir con su ayuda que los modelos evolutivos más complejos no sólo son posibles sino positivamente necesarios si estamos realmente dispuestos a explicar lo que el darwinismo siempre ha afirmado que quiere explicar. La idea de Darwin de que la evolución es un proceso algorítmico está ahora llegando a ser una familia de hipótesis aún más enriquecida, experimentando su propia

explosión de población gracias a la apertura de nuevos ambientes en los que vivir.

En el campo de la inteligencia artificial una estrategia que produce beneficios es la de trabajar sobre versiones deliberadamente simplificadas de los fenómenos que nos interesan. Estas versiones son denominadas con humor *problemas de juguete*. ¡En el mundo del juguete de hojalata de la biología molecular podemos conseguir ver en acción las versiones más simples de los fenómenos darwinianos fundamentales, pero éstas son problemas de juguetes *reales*! Se puede sacar ventaja de la relativa simplicidad y limpieza de esta teoría darwiniana del más bajo nivel para introducir e ilustrar algunos de los temas cuya pista seguiremos en los próximos capítulos a través de los niveles más elevados de la teoría de la evolución.

Los evolucionistas se han ayudado siempre a sí mismos, por ejemplo, con declaraciones acerca de la *optimalidad* y el crecimiento de la complejidad, y estas afirmaciones han sido reconocidas como serias hipersimplificaciones, en el mejor de los casos, por los que las hacían y por sus críticos. En el mundo de la evolución molecular estas apologías no son necesarias. Cuando Eigen habla de «optimalidad» presenta una clara definición de lo que quiere decir con este término, dispone de medidas experimentales para apoyarla y mantenerla en el terreno de la honestidad científica. Sus *paisajes adaptativos* y sus mediciones de los resultados no son subjetivas ni *ad hoc*. La complejidad molecular puede medirse por medio de varios métodos objetivos que se corroboran mutuamente; cuando Eigen utiliza el término «algoritmo» no lo hace en absoluto usando una licencia poética. Cuando, por ejemplo, imaginamos una enzima que corrige pruebas moviéndose a lo largo de un par de filamentos de ADN, revisando, reparando y haciendo copias para trasladarse después a otro escalón y repetir el proceso, es difícil que podamos poner en duda que estamos observando un autómata microscópico en pleno trabajo, y las mejores simulaciones se corresponden con los hechos observados tan ajustadamente que podemos estar seguros de que no hay duendes que los ayudan ni ganchos celestes ocultos por estos barrios. En el mundo de las moléculas, la aplicación del pensamiento darwiniano es particularmente pura y sin adulteración. Es natural que cuando adoptamos esta posición ventajosa pueda parecer asombroso que la teoría de Darwin, que funciona a la perfección en las moléculas, se aplique con todas sus consecuencias a unos desmañados conglomerados de células —en comparación, de tamaño galáctico— como son las aves, las orquídeas y los mamíferos. (Si no esperamos que la tabla periódica nos informe sobre corporaciones o naciones, ¿porqué debemos

esperar que la teoría evolucionista de Darwin funcione sobre tales complejidades como los ecosistemas o los linajes de mamíferos?).

En la biología *macroscópica* —la biología de los organismos con tamaños más comunes como hormigas, elefantes y secuoyas— todo es menos nítido. Las mutaciones y la selección sólo se suelen inferir de modo indirecto e imperfecto, gracias a una mareante colección de complicaciones circunstanciales. En el mundo molecular, los eventos de mutación y selección pueden ser directamente medidos y manipulados, y el tiempo de generación para los virus es tan corto que los grandes efectos darwinianos pueden ser estudiados. Por ejemplo, la espantosa capacidad de los virus tóxicos para mutar en un combate mortal con la moderna medicina es la que espolea y subvenciona la mayor parte de esta investigación. (¡El virus del sida ha experimentado tantas mutaciones en la última década que su historia en este período muestra más diversidad genética —medida en revisiones de codones — que la que se encuentra en toda la historia de la evolución de los primates!).

La investigación de Eigen y cientos de otros tiene aplicaciones prácticas bien definidas para todos nosotros. Es apropiado observar que este importante trabajo es un ejemplo de darwinismo triunfante, reduccionismo triunfante, mecanicismo triunfante y materialismo triunfante. Sin embargo, también es la cosa más alejada de un reduccionismo avaricioso. Es una grandiosa cascada de niveles sobre niveles, con nuevos principios de explicación, nuevos fenómenos que aparecen en cada nivel, siempre poniendo de manifiesto que la esperanza demasiado indulgente de explicarlo «todo» al nivel más bajo está mal encaminada. He aquí el resumen del propio Eigen, escrito en términos que deben ser compatibles con los críticos del reduccionismo más ardientes:

> La selección es como un demonio especialmente sutil que ha operado en los diferentes escalones para llegar a la vida, y opera hoy a diferentes niveles de la vida, con un conjunto de habilidades extraordinariamente original. Por encima de todo, es extraordinariamente activo, dirigido por un mecanismo interno de *feedback* que busca de un modo muy discriminative el mejor camino para conseguir una realización óptima, no porque posea una tendencia inherente hacia un objetivo predestinado, sino simplemente en virtud de su inherente mecanismo no lineal que le da la apariencia de dirigirse a un objetivo (Eigen 1992:123).

3. *Función y especificación*

En el mundo de las macromoléculas la forma es destino. Una secuencia unidimensional de aminoácidos (o de los codones de nucleótidos que los codifican) determina la identidad de una proteína, pero la secuencia tan sólo

limita parcialmente la vía en la que el filamento de esta proteína unidimensional se pliega sobre sí mismo. Se retuerce típicamente en una de las muchas formas posibles, a modo de una maraña conformada idiosíncráticamente, que es la que el tipo de su secuencia siempre prefiere. Esta forma tridimensional es la fuente de su poder, de su capacidad como agente catalítico, por ejemplo, como constructor de estructuras, combatiente de antígenos o regulador del desarrollo. Es una máquina, y lo que cumple es una función muy estricta derivada de la forma de sus partes constituyentes. Su forma global tridimensional es bastante menos importante, desde el punto de vista funcional, que la secuencia unidimensional que es responsable de esta función. Por ejemplo, la importante proteína lisozima es una máquina molecular de una forma particular, producida en muy diferentes versiones; más de cien diferentes secuencias de aminoácidos han sido encontradas en la naturaleza que se pliegan de la misma forma funcional y, naturalmente, las diferencias en las secuencias de estos aminoácidos pueden ser utilizadas como claves «filológicas» para recrear la historia evolutiva de la producción y del uso de la lisozima.

He aquí un rompecabezas, detectado por primera vez por Walter Elsasser [1958, 1966] pero definitivamente solucionado por Jacques Monod [1971]. Considerado de un modo muy abstracto, el hecho de que un código unidimensional pueda servir «para» una estructura tridimensional demuestra que se añade alguna información. Es decir, un *valor* es añadido. Los aminoácidos individualmente tienen valor (por su contribución a la habilidad funcional de una proteína), no precisamente en virtud de su localización en la secuencia unidimensional que forma el filamento, sino gracias a su localización en el espacio tridimensional, una vez que el filamento se ha plegado.

De este modo, hay una aparente contradicción entre la afirmación de que el genoma «define enteramente» la función de una proteína y el hecho de que esta función esté ligada a una estructura tridimensional cuyo contenido en datos es más rico que la contribución aportada por el genoma a la estructura (Monod 1971:94).

Como Küppers (1990:120) señaló, la solución de Monod es directa: «La aparentemente irreducible, o excesiva, información está contenida en las específicas condiciones del medio ambiente de la proteína, y sólo conjuntamente con estas condiciones la información genética puede determinar, sin ambigüedades, la estructura y, de este modo, la función de la molécula proteica». Monod lo ha expresado de esta manera:

De todas las estructuras posibles sólo una se realiza realmente. En consecuencia, las condiciones iniciales entran entre los ítems de información finalmente incluidos dentro de la estructura. Sin especificarlo, contribuyen a la realización de una única forma eliminando todas las estructuras alternativas, en este sentido proponiendo —o más bien imponiendo— una interpretación inequívoca de un mensaje potencialmente equívoco (Monod 1971:94)[64].

¿Qué significa esto? Sorprendentemente no significa que el lenguaje del ADN y los «lectores» de este lenguaje tengan que evolucionar conjuntamente; ninguno puede trabajar en solitario. Cuando los deconstruccionistas dicen que el lector aporta algo al texto están diciendo algo que se aplica con seguridad tanto al ADN como a la poesía; ese algo que el lector aporta puede ser caracterizado, de forma general y abstracta, como información, y tan sólo la combinación de información proporcionada por el código y del medio ambiente que lee el código son suficientes para crear un organismo[65]. Como hemos subrayado en el capítulo 5, algunos críticos han considerado este hecho como si fuera una refutación del «gencentrismo», la doctrina que postula que el ADN es la única información almacenada por herencia, aunque esta idea fue una hábil hipersimplificación. Aunque a las bibliotecas se las considera habitualmente como almacenes de información, en realidad son sólo las bibliotecas-*plus*-lectores las que preservan y almacenan la información. Dado que las bibliotecas no han contenido entre sus volúmenes —hasta ahora, a cualquier nivel— la información necesaria para crear más lectores, su capacidad para almacenar información ha dependido de *otro* sistema de almacenamiento de información, el sistema genético humano, del cual el ADN es el medio principal. Cuando aplicamos el mismo razonamiento al propio ADN vemos que, además, necesita un continuo aporte de «lectores» que no es, en sí mismo, enteramente específico. ¿De dónde llega el resto de la información para especificar a estos lectores? La breve respuesta es que llega de sus continuidades con su medio ambiente: la persistencia en el medio ambiente de los materiales necesarios (y de fragmentos ya parcialmente construidos o prefabricados) y de las condiciones en las cuales estos materiales pueden ser aprovechados. Cada vez que nuestro estropajo se seca entre dos usos, rompemos la cadena de la continuidad ambiental (por ejemplo, mucha humedad), que es parte del trasfondo de información presupuesta por el ADN de la bacteria en el estropajo, cuya destrucción buscamos.

Vemos aquí un caso especial de un principio muy general: cualquier estructura *funcionante* es vehículo de información *implícita* acerca del medio ambiente en el que su función «trabaja». Las alas de una gaviota encarnan magníficamente los principios del diseño aerodinámico, lo que también implica que la criatura que posee estas alas está muy bien adaptada para volar

en un medio que tenga la densidad específica y la viscosidad de la atmósfera dentro de unos miles de metros, poco más o menos, de la superficie de la Tierra. Recordemos el ejemplo del capítulo 5, que consistía en enviar la partitura de la Quinta Sinfonía de Beethoven a los «marcianos». Supongamos que hemos preservado cuidadosamente el cuerpo de una gaviota y lo enviamos al espacio (sin que le acompañe una explicación) para que sea descubierto por dichos marcianos. Si ellos establecen la fundamental hipótesis de que las alas eran funcionales y que su función era volar (lo cual, pienso, puede no ser tan obvio para los marcianos como para nosotros, que hemos visto cómo lo hacían las gaviotas) pueden utilizar esta conjetura para «leer» la información implícita acerca de un medio ambiente para el cual estas alas habían sido bien diseñadas. Supongamos que los marcianos se preguntan a sí mismos de qué modo toda esta teoría aerodinámica está implícita en la estructura o, en otras palabras: ¿cómo toda esa información se aplica a las alas? La respuesta *debe ser*: por medio de una interacción entre el medio ambiente y los ancestros de la gaviota. (Dawkins 1983a ha investigado estos temas con más detalles).

Los mismos principios se aplican al nivel más básico, donde la función es en sí misma especificación, *la función de las que todas las otras funciones dependen*. Cuando nos preguntamos, igual que Monod, cómo queda fijada la forma tridimensional de las proteínas, dado que la información en el genoma debe especificarla de manera insuficiente, vemos que sólo puede explicarlo una poda de lo no funcional (o de lo menos funcional). De este modo, la adquisición de una forma particular por una molécula implica una mezcla de accidente histórico por un lado y del «descubrimiento» de verdades importantes por otro.

Desde el comienzo, el proceso de diseño de las «máquinas» moleculares exhibe estas dos características de la ingeniería humana. Eigen [1992:34] propone un buen ejemplo de este hecho en su reflexión sobre la estructura del código del ADN. «Uno puede preguntar por qué la naturaleza ha usado cuatro símbolos, cuando podía haberlo hecho bien con dos». ¿Por qué? Obsérvese como, natural e inevitablemente, un «por qué» se suscita en esta cuestión, y cómo también se solicita una respuesta de «ingeniería». La respuesta puede ser, alternativamente, que no hay razón para ello —es un puro y simple accidente histórico— o *hay* una razón: había o hay una condición que hace que ésta sea la vía *correcta* o la *mejor* vía para diseñar el código genético, dadas las condiciones que ha obtenido[66].

Todos los hechos más profundos del diseño molecular deben considerarse desde la perspectiva de la ingeniería. Por un lado, el hecho de que la forma de las macromoléculas se presente en dos categorías básicas: simétrica y asimétrica o *chiral* (en las versiones mano izquierda y mano derecha). Esta es la razón por la que muchas deben ser simétricas:

La ventaja selectiva en un complejo simétrico es disfrutada por todas las subunidades, mientras que en los complejos asimétricos la ventaja es solamente efectiva en las subunidades en las que se produce la mutación. Por esta razón encontramos tantas estructuras simétricas en biología, «a causa de que estas estructuras son capaces de hacer el uso más efectivo de su ventaja y de este modo —*a posteriori*— ganar la competición de la selección; sin embargo, esto no es así porque la simetría sea —*a priori*— un requerimiento indispensable para el cumplimiento de un propósito funcional». (Küppers 1990:119, incorporando una cita de Eigen y Winkler-Oswatitsch 1975).

Pero ¿qué podemos decir de las formas asimétricas? ¿Hay una razón por la que estas formas deban ser así; es decir, zurdas y no diestras? Probablemente no, pero «aun si no hubiera *a priori* una explicación física para la decisión, si fuera justamente una ligera fluctuación la que diese lugar a una u otra posibilidad equivalente a una ventaja momentánea, el carácter de autorreforzamiento de proceso selectivo convertiría la decisión aleatoria en una ruptura de la simetría mayor y permanente. La causa sería puramente “histórica”» (Eigen 1992:35)**[67]**.

En nuestra parte del universo, la asimetría compartida de las moléculas orgánicas sería otro puro fenómeno QWERTY, o lo que Crick [1968] ha denominado un *accidente congelado*. Pero aun en el caso de que se tratase de un fenómeno QWERTY, si esta condición es precisamente correcta y las oportunidades, y en consecuencia las presiones, son bastante grandes, las tablas pueden cambiarse y establecerse un nuevo estándar. Esto es en apariencia lo que justamente sucede cuando el lenguaje del ADN desplazó al lenguaje del ARN como la *lingua franca* de las codificaciones y descodificaciones en los organismos complejos. Las *razones* para esta preferencia son claras: el lenguaje del ADN, al constar de dos filamentos, permite un sistema de enzimas para la corrección de errores o la lectura de pruebas, las cuales pueden reparar los errores de copia en un filamento comparándolo con su par. Esto hace posible la creación de genomas más largos y más complicados (Eigen 1992:36).

Este razonamiento *no* da lugar a la conclusión de que la doble hélice del ADN debe desarrollarse, ya que la madre naturaleza no tiene intención de crear por adelantado la vida multicelular. Lo que revela es que si el ADN con doble hélice empezó a desarrollarse, abrió oportunidades que dependían de él.

En consecuencia, llegó a ser una necesidad para aquellos ejemplares en el espacio de todas las formas de vida posibles que se benefician de este hecho, y si aquellas formas de vida prevalecen sobre aquellas otras que no se benefician de este hecho, se genera un apoyo retroactivo de la *raison d'être* del lenguaje del ADN. Este es el camino por el que la evolución siempre descubre razones, por apoyo retroactivo.

4. *El pecado original y el nacimiento del significado*

> ¿El camino hacia la sabiduría?
> Bien, es muy sencillo y simple de expresar:
> equivocarse y equivocarse y equivocarse una y otra vez pero cada vez menos y menos y menos.
>
> *Piet Hein*

> La solución al problema de la vida se observa en el desvanecimiento de este problema.
>
> Ludwig Wittgenstein,
> *Tractates Logico-philosophicus, 6*

Erase una vez una época en que no había mente, ni significado, ni error, ni función, ni razones, ni vida. Ahora, todas esas admirables cosas existen. Tiene que ser posible contar la historia de cómo llegó a existir todo esto y esta historia debe pasar, mediante sutiles incrementos, desde los elementos que carecen de estas maravillosas propiedades, a los elementos que manifiestamente las tienen. Allí tendría que haber istmos de intermediarios, dudosos, controvertidos o, simplemente, inclasificables. Todas estas maravillosas propiedades deben haber llegado a existir gradualmente, por etapas que son apenas perceptibles, incluso en una *visión retrospectiva*.

Recuérdese que en el capítulo anterior parecía obvio, quizás una verdad de la lógica, que debería haber existido o un primer ser vivo o una infinita regresión de seres vivos. Naturalmente, no era ni una ni otra alternativa y la solución darwiniana estándar, con la que nos volveremos a encontrar una y otra vez, era ésta: en su lugar describimos una regresión finita, en la que la maravillosa propiedad buscada —en este caso la vida— se consiguió a través de leves, quizás incluso imperceptibles, rectificaciones o incrementos.

He aquí la forma más general del esquema de la explicación darwiniana. La tarea de salir de los primeros tiempos —cuando no había ninguna *x*—

hasta alcanzar el tiempo posterior —cuando había grandes cantidades de *x*— se completa con una serie finita de pasos en los cuales llega a ser cada vez menos claro que «realmente aquí todavía no hay ninguna *x*», a través de una serie de pasos «discutibles», hasta que, eventualmente, nos encontramos a nosotros mismos en pasos en donde es bastante obvio que «por supuesto hay una *x*, muchas *x*». Nunca trazaremos líneas.

Obsérvese lo que sucede en el caso particular del origen de la vida si tratamos de trazar la línea. Hay un montón de verdades —sin duda muchos de sus detalles son desconocidos para nosotros, cualquiera de las cuales podríamos identificar, *en principio*, si quisiéramos, como la verdad que confirma la identidad de «Adan el Protobacterium». Sin embargo, podemos precisar, como queramos, las condiciones para ser el primer ser vivo, pero cuando al utilizar la máquina del tiempo volvemos atrás para ser testigos del momento, encontramos que «Adan el Protobacterium», no importa cómo lo hayamos definido, es probablemente tan indistinguible como la Eva Mitocondrial. Sabemos, sobre bases lógicas, que hubo al menos un comienzo que tiene en nosotros su continuación, pero que hubo probablemente muchos falsos comienzos que no difieren, en último término, *en ninguna interesante peculiaridad* de aquél que inició la serie ganadora. El título de Adán, recordémoslo, es un honor retrospectivo y cometeríamos un error fundamental de razonamiento si preguntásemos: *¿en virtud de qué diferencia esencial éste es el comienzo de la vida?* No es necesario que exista diferencia entre Adán y Badán, un duplicado de Adán átomo por átomo que no es el comienzo de nada notable. Este no es un *problema* de la teoría de Darwin; es una fuente de su poder. Como Küppers ha señalado: «El hecho de que nosotros no estemos en condiciones de presentar obviamente una definición comprensiva del fenómeno de la "vida" no dice nada en contra, sino a favor de la posibilidad de una completa explicación física del fenómeno de la vida» (Küppers 1990:133).

A este mismo predicamento gratuito se enfrenta cualquiera que, perdida la esperanza de definir algo tan complicado como la vida, decide definir las nociones de *función* o *teleología* aparentemente más simples. ¿En qué momento exacto hace su aparición la función? ¿Tenían función o poderes causales los primeros nucleótidos? ¿Exhibieron los cristales de arcilla de Cairns-Smith propiedades teleológicas genuinas o algo así *como* si fueran propiedades teleológicas? ¿En el mundo del juego de la vida desempeñan la función de locomoción los *deslizadores* o simplemente se mueven? No hay mucha diferencia en cómo legislemos la respuesta; el mundo de los

mecanismos que funcionan debe comenzar con mecanismos que «estén a caballo de la línea fronteriza» y, sin embargo, bastante más atrás de donde coloquemos dicha línea, habrá precursores que difieran en peculiaridades no esenciales, de los precursores consagrados[68].

Nada que pueda ser lo suficientemente complicado como para ser interesante podría tener una esencia[69]. Esta cuestión antiesencialista fue reconocida por Darwin como un acompañamiento epistemológico o metafísico a su ciencia, verdaderamente revolucionario; no debe sorprendernos lo difícil que resulta aceptar tal premisa. Desde que Sócrates enseñó a Platón —y al resto de nosotros— cómo jugar al juego de las interrogaciones sobre cuestiones necesarias y suficientes, hemos considerado la tarea de «definir nuestros términos» como un apropiado preámbulo para todas las investigaciones serias y esto nos ha enfrentado a interminables asaltos de los «traficantes de esencias»[70]. *Queremos* trazar líneas de fronteras; a menudo *necesitamos* trazar líneas, porque así podemos dar por acabadas o prevenir a tiempo exploraciones estériles. Nuestros sistemas de percepción han sido genéticamente diseñados para forzar a los candidatos que se encuentran con un pie a cada lado de la línea fronteriza, en una posición ambigua, a ser atribuidos en una clasificación u otra (Jackendoff 1993), lo cual es un buen truco, pero no un movimiento forzoso. Darwin demostró que la evolución no necesita lo que nosotros necesitamos; el mundo real puede seguir adelante perfectamente con las divergencias *de facto* que emergen a lo largo del tiempo, dejando muchos vacíos entre agolpamientos de realidad.

Hemos dirigido brevemente nuestra atención a una cuestión bastante importante de este característico esquema explicativo darwiniano, y conviene hacer una pausa para confirmar el efecto. A través del microscopio de la biología molecular hemos sido testigos del nacimiento, en la primeras macromoléculas que tenían la suficiente complejidad como para *hacer cosas*, de la *capacidad de actuar u obrar* como agentes. No es esta una capacidad de actuar compleja, una acción intencional *auténtica*, con la representación de razones, deliberación, reflexión y decisión consciente, aunque este es el único terreno posible en el que las semillas de la acción intencional pueden crecer. Hay algo extraño y vagamente repelente en esta *casi capacidad de acción* que hemos descubierto a este nivel, en todo este trajín intencional y apresurado, aunque *todavía no hay nadie en casa*. Las máquinas moleculares realizan sus admirables habilidades, es obvio que exquisitamente diseñadas, y tan obvio como que no se sabe más que antes de lo que estas máquinas están haciendo.

Veamos este relato acerca de la actividad de un fago de ARN, un virus replicante:

> En primer lugar, el virus necesita de un material en el que pueda empaquetar y proteger su propia información. En segundo lugar, necesita un medio para introducir su información en la célula huésped. En tercer lugar, requiere un mecanismo para la replicación específica de su información en presencia de un vasto exceso de ARN de la célula huésped. Finalmente, debe disponerse para la proliferación de su información, un proceso que suele conducir a la destrucción de la célula huésped (…) El virus consigue entonces la célula para llevar a cabo su replicación: su única contribución es un factor proteico, especialmente adaptado para el ARN viral. Esta enzima no se activa hasta que aparece una contraseña en el ARN viral. Cuando aparece esta contraseña reproduce el ARN viral con gran eficiencia, mientras que ignora el número mucho mayor de moléculas de ARN de la célula huésped. En consecuencia, la célula es rápidamente inundada con el ARN viral. Éste es empaquetado en la cubierta de la proteína del virus, que es también sintetizada en grandes cantidades, y finalmente la célula huésped estalla y suelta una numerosísima progenie formada por partículas virales. Todo esto es un programa que se desarrolla automáticamente y de acuerdo con lo ensayado, en sus más pequeños detalles (Eigen 1992:40).

Gusten o no gusten, fenómenos como éste ponen al descubierto la esencia del vigor de la idea darwiniana. Un pequeño fragmento de maquinaria molecular, impersonal, irreflexiva, robótica, insensata es, en último término, la base de roda la acción y, en consecuencia, del significado y, en consecuencia, de la conciencia en el universo.

Desde el comienzo, el coste de *hacer algo* corre paralelo al riesgo de hacerlo *erróneamente*, de cometer una equivocación. Nuestro eslogan puede ser: no se consigue nada sin error. El primer error que se cometió era un error tipográfico, un error de copia que entonces tuvo la oportunidad de crear una nueva tarea en el medio ambiente (o *paisaje adaptativo*) con un nuevo criterio de lo que es correcto y de lo que es erróneo, de lo que es mejor y de lo que es peor. Un error de copia «cuenta» aquí como un único error porque cometerlo tiene un coste: terminación de la línea reproductiva en el peor de los casos o disminución de la capacidad de replicación. Todas éstas son cuestiones objetivas, diferencias que están ahí tanto si las observamos o nos preocupamos de ellas, como si no, pero que aportan una nueva perspectiva. Hasta ese momento no existía ninguna oportunidad para el error. Sin embargo, las cosas no marchaban ni bien ni mal. Hasta ese momento no existía un método estable y predictivo para poder optar por la adopción de la perspectiva desde la que los errores podían ser percibidos, y cada error que cualquiera o cualquier cosa ha cometido depende de este original proceso que produce errores. En efecto, hay una fuerte presión de selección que hace que el proceso de la copia genética se haga con la más alta fidelidad posible, minimizando la posibilidad del error. Afortunadamente, no se puede alcanzar la perfección completa, porque si fuera así, la evolución se detendría. Éste es

el pecado original en una versión científicamente respetable. Como en la versión cristiana, trata de explicar algo: la emergencia de un nuevo nivel de fenómenos con características especiales (lectores de significados en una versión, pecadores en la otra). Al contrario que en la versión cristiana, nos aporta una explicación que tiene sentido; no se proclama a sí misma como un hecho misterioso que uno debe aceptar mediante la fe, y sus implicaciones pueden ser sometidas a comprobación.

Cabe señalar que uno de los primeros frutos de la perspectiva según la cual el error es percibido es una aclaración del concepto de especie. Cuando consideramos todos los textos genómicos actuales que han sido creados en el repetido proceso de la copia una y otra vez —con mutaciones ocasionales— nada cuenta intrínsecamente como una versión canónica de algo. Es decir, mientras podemos identificar mutaciones por simple comparación entre la secuencia «anterior» y la «posterior», no hay un modo intrínseco de decir cuál de los *errores* tipográficamente incorrectos puede ser considerado fructífero, como una *mejora* editorial**[71]**. La mayoría de las mutaciones son lo que los ingenieros llamarían variaciones *sin importancia* que no provocan diferencias perceptibles en cuanto a la viabilidad, pero como la selección se cobra su peaje gradualmente, la mejor versión comienza a acumularse. Sólo en relación a un *tipo salvaje* —que se comporta como el centro de gravedad de tal acumulación— podemos identificar una versión concreta como la versión de un error, e incluso cabe la posibilidad, remota en la práctica pero omnipresente en principio, de que lo que consideramos un error desde la perspectiva de un tipo salvaje es, en realidad, una brillante mejora, considerado desde la perspectiva de un nuevo *tipo salvaje en marcha*. Y cuando emergen nuevos tipos salvajes como los focos o las cimas de *paisajes adaptativos*, la dirección de la presión de la *corrección* del error puede ser revertida en una concreta vecindad del espacio de diseño. Una vez que una familia particular de textos similares deja de ser sujeto de *corrección* en relación a una norma retirada o caducada, esta familia es libre para desplazarse por la atractiva cuenca de una nueva norma**[72]**. El aislamiento reproductivo es, pues, tanto la causa como el efecto del amontonamiento en el espacio fenotípico. Donde quiera que haya regímenes de competencia en la corrección de errores, un régimen ganará al otro y en consecuencia el istmo entre los competidores tenderá a disolverse, dejando un espacio vacío ente las zonas ocupadas del espacio de diseño. Así, del mismo modo que las normas de pronunciación y el uso de las palabras refuerza el agrupamiento en comunidades lingüísticas (una importante cuestión teórica planteada por

Quine en 1960 en su discusión sobre el error y la emergencia de normas en el lenguaje), las normas de la expresión genómica son la base fundamental de la especiación.

A través del mismo nivel molecular microscópico observamos el nacimiento del *significado* en la adquisición de una «semántica» por parte de las secuencias de nucleótidos, las cuales, en un principio, son meros objetos sintácticos. Este es un paso crucial en la campaña darwiniana para derrotar a la concepción de John Locke del cosmos como una primera mente. Los filósofos suelen estar de acuerdo, por una buena razón, en que significado y mente no pueden ser nunca separados, que no puede haber significado si no hay mente, o que donde haya mente no haya significado. La *intencionalidad* es el término técnico de los filósofos para designar el significado; es el *acerca de qué* lo que puede relacionar una cosa con otra; un nombre para el que lo lleva, una llamada de alarma para el peligro que acecha, una palabra para su referente, un pensamiento para su objeto[73]. En el universo sólo algunas cosas manifiestan intencionalidad. Un libro o un cuadro pueden tratar de una montaña, pero una montaña en sí misma no trata de nada. Un mapa, un signo, un sueño o una canción pueden hablar de París, pero París en sí mismo no habla de nada. Los filósofos consideran la intencionalidad como la *marca* de lo mental. ¿De dónde viene la intencionalidad? De las mentes, naturalmente.

Sin embargo, la idea, perfectamente válida en sí misma, se convierte en una fuente de misterio y confusión cuando se utiliza como un principio metafísico, más que como un hecho de la historia natural reciente. Aristóteles llamaba a Dios el Motor Inmóvil, la fuente de todo movimiento en el universo, y la versión de Locke de la doctrina aristotélica, como hemos visto, identifica a Dios con la Mente, convirtiendo al Motor Inmóvil en lo que él denomina como «Unmeant Meaner», es decir, en la Fuente de Toda Intencionalidad («Meaner») sin que la suya le venga de otro («Unmeant»). Locke se impuso la tarea de probar deductivamente lo que la tradición ya había dado por obvio: la intencionalidad original surge de la mente de Dios; nosotros somos criaturas de Dios y derivamos nuestra intencionalidad de Él.

Darwin puso esta doctrina patas arriba: la intencionalidad no viene de arriba; se infiltra desde abajo hacia arriba, a partir de un proceso algorítmico inicial, insensato y sin objetivo, que gradualmente adquiere significado e inteligencia cuando se desarrolla. Y, siguiendo perfectamente el patrón de todo el pensamiento darwiniano, vemos que el primer significado no es un significado que haya desplegado complemente sus alas; falla en poner de manifiesto todas las propiedades «esenciales» del significado real (pensemos

lo que pensemos acerca de lo que estas propiedades son). Es simplemente un *casi significado* o una semisemántica. Es lo que John Searle [1980, 1985, 1992] ha menospreciado denominándolo como casi *intencionalidad*, como lo opuesto a lo que él llama *intencionalidad original*. Pero debemos empezar en algún lugar y el hecho de que el primer paso en la dirección correcta sea escasamente discernible como un paso hacia la significación es, en último término, lo que debemos esperar.

Hay dos caminos que conducen a la *intencionalidad*. La vía darwiniana es diacrónica o histórica y tiene que ver con el aumento gradual, durante miles de millones de años de los tipos de diseño —en cuanto a funcionalidad y a propósito— que puedan soportar una interpretación intencional de las actividades de los organismos (las *cosas* que *hacen* los *agentes*). Antes de que la *intencionalidad* pueda desplegar completamente sus ya emplumadas alas, debe pasar por un difícil y feo período de pseudointencionalidad *sin plumas*. El camino sincrónico es el de la Inteligencia Artificial: en un organismo con *intencionalidad* genuina —como nosotros— hay ahora mismo muchas partes y algunas de estas partes exhiben un cierto tipo de *semiintencionalidad*, o simplemente casi *intencionalidad*, o *pseudointencionalidad* —llámese como se quiera—, y nuestra propia, genuina y muy desplegada intencionalidad es, de hecho, el producto (sin posteriores ingredientes milagrosos) de las actividades de todos los *bits* insensatos o medio insensatos que terminan por concebirnos (ésta es la tesis central defendida en Dennett 1987b, 1991a). Una mente es eso; no una máquina milagrosa, sino una enorme, mediodiseñada, autorrediseñada amalgama de máquinas más pequeñas, cada una de las cuales tiene un diseño con su propia historia y desempeña su propio papel en la «economía del alma». (Platón estaba en lo cierto, como suele ocurrir, cuando observó una profunda analogía entre una república y una persona; aunque naturalmente tenía una visión demasiado simple de lo que esto podría significar).

Existe una profunda afinidad entre los caminos, el sincrónico y el diacrónico, que conducen a la intencionalidad. Un modo de dramatizar este asunto es parodiar un antiguo sentimiento anti-Darwin: el tío del mono. ¿Estaría dispuesto a que su hija se casara con un robot? Bien, ¡si Darwin estaba en lo cierto, su bisbisbis… abuela *era* un robot! De hecho, era un *macro*. Esta es la inevitable conclusión de los capítulos anteriores. No sólo descendemos de macros, sino que estamos compuestos de macros. Nuestras moléculas de hemoglobina, nuestros anticuerpos, nuestras neuronas, la maquinaria de nuestro reflejo vestíbulo-ocular; en cada nivel de análisis

encontramos maquinaria que silenciosamente hace un trabajo admirable y elegantemente diseñado. Quizás hayamos dejado de temblar ante la visión científica de los virus y de las bacterias que con diligencia e insensatez ejecutan sus proyectos subversivos; horribles enanos autómatas que realizan sus malévolas acciones. Pero no debemos creer que estaremos más cómodos si pensamos que (*ellos*) son invasores alienígenas, es decir, distintos de los tejidos más sociables que nos han hecho a nosotros. Estamos hechos de los mismos tipos de autómatas que nos invaden; ninguna aureola de *élan vital* distingue a nuestros anticuerpos de los antígenos contra los que combaten; simplemente pertenecen a nuestro club, así que los combaten en nuestro nombre.

¿Quiere esto decir que si reunimos un número suficiente de esos homúnculos silenciosos formarían una persona real y consciente? Los darvinianos dicen que no hay otra manera de hacerlo. Ahora bien, del hecho de que seamos descendientes de *robots* no se deriva que nosotros seamos robots. Después de todo nosotros somos también descendientes directos de algún pez y no somos peces; somos descendientes directos de alguna bacteria y no somos bacterias. Pero a menos que el dualismo o el vitalismo sean verdad (en cuyo caso tendríamos en nosotros algún secreto ingrediente extra), estamos *hechos* de robots, o lo que viene ser lo mismo, somos una colección de trillones de máquinas macromoleculares. Y todas ellas descienden de los macros originales. De este modo, algo hecho de robots puede exhibir una conciencia genuina, o una genuina intencionalidad, porque nosotros funcionamos si todo funciona.

No es de extrañar, pues, que exista tanto antagonismo contra el pensamiento darwiniano como contra la Inteligencia Artificial. Ambos asestan, conjuntamente, un golpe fundamental al último refugio al que se había retirado la gente frente a la revolución copernicana: a la mente como un sanctasanctórum que la ciencia no puede alcanzar (véase Mazzlish 1993). Es una larga y amplia carretera que va desde las moléculas hasta las mentes, con muchos espectáculos divertidos a lo largo del camino, y sólo nos demoraremos ante los más interesantes en los siguientes capítulos. Pero ahora ha llegado el momento de contemplar más de cerca los comienzos darwinianos de la Inteligencia Artificial.

5. *El ordenador que aprendió a jugar al ajedrez*

Alan Turing y John von Neumann han sido dos de los más grandes científicos de este siglo. De ellos se puede decir que inventaron el ordenador, y su invento ha llegado a ser reconocido como un triunfo de la ingeniería y un vehículo intelectual para explorar los territorios más abstractos de la ciencia pura. Ambos pensadores fueron, al mismo tiempo, imponentes teóricos y profundos prácticos, paradigmas de un estilo intelectual que ha ido desempeñando un papel creciente en la ciencia desde la Segunda Guerra Mundial. Además de crear el ordenador, tanto Turing como Von Neumann aportaron contribuciones fundamentales a la biología teórica. Von Neumann, como ya hemos señalado, aplicó su brillante mente al problema abstracto de la autorreplicación, y Turing [1952] realizó trabajos pioneros en uno de los problemas teóricos más básicos de la embriología o morfogénesis: ¿cómo puede la compleja topología —es decir, la forma— de un organismo surgir de la simple topología de una única célula fertilizada desde la que crece? El proceso comienza, como cualquier estudiante de bachillerato sabe, con una división completamente simétrica (como ha escrito François Jacob, el sueño de toda célula es convertirse en dos células). Dos células llegan a ser cuatro, y cuatro llegan a ser ocho y ocho llegan a ser dieciséis; ¿cómo se inicia la formación del corazón, del hígado, de las extremidades y del cerebro en tales condiciones[74]? Turing adivinó la similitud entre estos problemas a nivel molecular, y el problema de cómo un poeta escribe un soneto, por lo que desde los primeros días de los ordenadores la ambición de aquellos que compartían la intuición de Turing ha sido utilizar su admirable máquina para explorar los misterios del pensamiento[75].

Turing publicó en 1950 su profético ensayo «Computing Machinery and Intelligence» en la revista de filosofía *Mind*, seguramente uno de los artículos citados con mayor frecuencia en dicha revista. Cuando lo escribió, no existían programas de inteligencia artificial; realmente sólo existían dos ordenadores en el mundo, pero al cabo de pocos años habría suficientes máquinas funcionando las veinticuatro horas del día, de modo que Arthur Samuel, un científico que investigaba para la IBM, pudo ocupar parte de su tiempo de ocio a últimas horas de la noche en uno de los primeros ordenadores gigantes con las actividades de un programa que es tan buen candidato como cualquier otro para el título de Adán-IA (Inteligencia Artificial). El programa de Samuel jugaba a las damas y mejoraba su juego jugando contra sí mismo por las noches, rediseñándose a base de eliminar versiones que no habían funcionado bien en sus partidas nocturnas y sometiendo a prueba nuevas mutaciones que habían sido generadas de manera no inteligente. El ordenador

llegó a ser eventualmente mejor jugador de damas que el propio Samuel, aportándonos uno de los primeros claros contraejemplos del mito, de algún modo histérico, que dice que «un ordenador puede realmente hacer tan sólo lo que el programador le ha dicho que haga». Desde nuestra perspectiva podemos comprobar que esta común pero errónea idea es sólo una expresión del presentimiento de Locke de que sólo la mente puede diseñar, una explotación del *ex nihilo nihil fit* que Darwin claramente había desacreditado. Por otra parte, la vía por la que el programa de Samuel trasciende a su creador fue un llamativo y clásico proceso de evolución darwiniana.

El legendario programa de Samuel no es sólo el progenitor de las especies intelectuales, la Inteligencia Artificial, sino también de sus más recientes progenies, como es la Vida Artificial (AL por «Artificial Life»). Aunque se trata de un viejo programa, poca gente conoce hoy sus extraordinarios detalles, muchos de los cuales merecen ser conocidos con más profundidad[76]. El primer programa de damas de Samuel fue escrito en 1952 para el IBM 701, pero la versión de aprendizaje no estuvo terminada hasta 1955 para un IBM 704; una versión posterior se produjo para el IBM 7090, en la que Samuel encontró algunas vías elegantes para codificar en una partida de damas cualquier situación dentro de cuatro «palabras» de treinta y seis bits y cualquier movimiento con una simple operación aritmética con aquellas palabras. (Comparado con los prodigiosos y derrochadores programas de hoy, que funcionan con megabytes, el programa básico de Samuel es microscópico en tamaño; un «genoma» de baja tecnología con poco más de seis mil líneas codificadas. En aquel momento Samuel tuvo que escribirlo en código máquina: esto sucedía antes de la introducción de los lenguajes para programar ordenadores). Una vez que resolvió el problema de representar el proceso básico de las reglas del juego, tuvo que hacer frente a la parte del problema verdaderamente difícil: conseguir que el programa hiciera la *evaluación* de los movimientos, de modo que pudiera *seleccionar* el mejor posible (o al menos uno de los mejores movimientos).

¿Cuál sería la evaluación correcta? Algunos juegos triviales como el *tic-tac-toe* tienen soluciones algorítmicas factibles. La victoria o la derrota de un jugador está garantizada y su mejor estrategia puede ser computada en un período de tiempo realista. El ajedrez no es un juego de estas características. Samuel (1964:72) puntualizó que el espacio de posibles juegos de damas contiene del orden de 10^{40} puntos de elección, «lo que significa que, a tres elecciones por milimicrosegundo, se tardarían 10^{21} siglos para considerarlos». Aunque los ordenadores actuales son millones de veces más rápidos que los

pesados armatostes de la época de Samuel, aún no pueden dar en la diana del problema utilizando el camino de la fuerza bruta como sería llevar a cabo una búsqueda exhaustiva. El espacio de la búsqueda es «vasto», por lo que el método de ésta debe ser heurístico, es decir, como el de un árbol que se ramifica siguiendo todos los movimientos posibles, el cual debe ser podado, de manera despiadada, por demonios miopes semiinteligentes que dirigen una exploración arriesgada y azarosa en una mínima subporción de la totalidad del espacio.

La búsqueda heurística es una de las ideas fundamentales de la Inteligencia Artificial. En este sentido, la tarea a realizar en el campo de la IA se puede definir como la creación y la investigación de algoritmos heurísticos. Sin embargo, dentro de la ciencia de la computación y de las matemáticas existe la tradición de *contrastar* los métodos heurísticos con los métodos algorítmicos: los métodos heurísticos son arriesgados y no garantizan la obtención de resultados, mientras que los algoritmos aportan una garantía. ¿Cómo resolver esta «contradicción»? En realidad, no existe contradicción. ¡Los algoritmos heurísticos son, como todos los algoritmos, procedimientos mecánicos garantizados para hacer lo que hacen, pero lo que hacen también es implicarse en una búsqueda arriesgada! No hay garantía de *encontrar* algo o, al menos, no hay garantía de encontrar, en el tiempo disponible, algo específico en lo que se había pensado. Pero, como los torneos de habilidades bien organizados, los buenos algoritmos heurísticos *tienden* a conseguir resultados muy interesantes y fiables en períodos razonables de tiempo. Son arriesgados, pero los buenos algoritmos heurísticos son, naturalmente, riesgos asumibles. Se puede apostar la vida por ellos (Dennett 1989b). El error de no tener en cuenta que los algoritmos pueden ser procedimientos heurísticos ha confundido a más de un crítico de la Inteligencia Artificial. En especial, ha confundido a Roger Penrose, cuyos puntos de vista serán analizados en el capítulo 15.

Samuel comprobó que la exploración del «vasto» espacio del juego de damas sólo podía ser factible si se hacía mediante un proceso que, arriesgadamente, fuera podando el árbol de la búsqueda; pero ¿cómo llevar a cabo la poda y escoger a los demonios que hicieran este trabajo? ¿Qué reglas fácilmente programables —en cuanto a la suspensión de la búsqueda en determinados momentos o a las funciones de evaluación— pueden tener una potencia superior al azar, de modo que el árbol de la búsqueda vaya creciendo en la dirección más sabia? Samuel perseguía el hallazgo de un buen método algorítmico de búsqueda. Procedía empíricamente comenzando por el diseño

de vías que permitieran mecanizar cualquier obvia regla práctica que se le ocurría. Lo hizo pensando con cuidado en lo que iba a hacer y aprendiendo de sus propios errores, por lo que el sistema debía tener una memoria en la que se almacenara la experiencia pasada. El método de *aprender memorizando* hizo que el prototipo del programa progresara mucho mediante el almacenamiento de miles de posiciones de partidas ya conocidas de las que se sabían sus resultados. Pero el *aprendizaje memoristico* no puede llevarnos demasiado lejos; el programa de Samuel se enfrentó en el proceso con una rápida disminución de la recuperación, cuando había almacenado cerca de un millón de palabras que describían la experiencia pasada y empezó a sobrecargarse con problemas de índices y de recuperación. Cuando se requería una ejecución de mayor nivel o más versátil, había que introducir una estrategia diferente en el diseño: la *generalización*.

En lugar de tratar de encontrar él mismo el procedimiento de búsqueda, Samuel trató de conseguir que lo encontrara el ordenador. Su objetivo era que el ordenador diseñara su propia función de evaluación, una fórmula matemática —un polinomio— que generara un número, positivo o negativo, para cada movimiento considerado, de modo que, en general, mientras más elevado fuera el número, mejor sería el movimiento. El polinomio tenía que estar fabricado por muchas piezas, cada una de ellas contribuyendo de forma positiva o negativa, multiplicado por uno u otro coeficiente y ajustado a otras circunstancias diferentes; sin embargo, Samuel no tenía ni idea de qué tipo de fabricación funcionaría bien. Confeccionó hasta treinta y ocho fragmentos a los que llamó *términos* y los introdujo en un fondo común. Algunos de estos términos eran intuitivamente valiosos, como aquellos que otorgaban puntos por aumento de la movilidad o por capturas potenciales de piezas, pero otros eran, en mayor o menor grado, muy poco convencionales, como el *término* DYKE definido así: «El parámetro está acreditado con I por cada línea de piezas pasivas que ocupe tres cuadrados adyacentes a una diagonal». En un determinado momento, dieciséis de los *términos* fueron colocados juntos en el «genoma funcionante» del polinomio activo y el resto quedó sin uso apropiado. A través de un conjunto de conjeturas inspiradas y de consonancias y chapuzas aún más inspiradas, Samuel diseñó reglas para la eliminación de la partida y encontró vías para mantener la «olla a presión», de modo que el proceso de *ensayo-error* fuera apropiado para generar buenas combinaciones de *términos* y coeficientes y reconocerlas en cuanto aparecían. El programa fue dividido en dos partes: *Alfa*, un pionero de rápidas mutaciones, y *Beta*, un oponente conservador que jugaba la versión que había

ganado el juego más reciente. *Alfa* generaliza su experiencia después de cada movimiento para ajustar los coeficientes en su evaluación polinómica y para reemplazar términos que no parecen importantes por nuevos parámetros extraídos de la lista de reserva (Samuel 1964:83).

Al principio se escogió una selección arbitraria de 16 términos y a todos los términos se les asignaron pesos iguales… Durante (las primeras partidas) fueron descartados y reemplazados un total de 29 términos diferentes, la mayoría en dos ocasiones diferentes… La calidad del juego fue extremadamente pobre. Durante las siguientes siete partidas hubo al menos ocho cambios en la cabecera de la lista que afectaron a cinco términos diferentes… La calidad del juego mejoró progresivamente pero la máquina aún jugaba bastante mal… Algunos buenos jugadores aficionados que jugaron con la máquina en este período (después de siete juegos más) coincidieron en que «la máquina era habilidosa pero batible» (Samuel 1964:89).

Samuel observó que, mientras que el aprendizaje en esta fase inicial era sorprendentemente rápido, también era «muy errático y no demasiado estable». Fue descubriendo que el espacio-problema que estaba siendo explorado era un paisaje adaptativo accidentado en el que un programa que utilizara la simple técnica de *escalar la colina* era propicio a caer en trampas, inestabilidades y asas obsesivas de las cuales el programa no podía recuperarse sin un golpecito de ayuda, o dos, propinado por su diseñador. El programa era capaz de reconocer en su sistema los «defectos» responsables de estas inestabilidades y corregirlos. El sistema final, aquel que batió a Nealey, era una amalgama del tipo Rube Goldberg de aprendizaje memorístico, *kludges*[77] y productos de autodiseño, inescrutables para el propio Samuel.

No es de extrañar que el programa de Samuel causara una tremenda sensación y estimulara enormemente a los primeros visionarios de la Inteligencia Artificial, pero el entusiasmo por tales algoritmos se desvaneció pronto. Mientras más se dedicaba la gente al intento de extender sus métodos a problemas más complejos —ajedrez, por ejemplo, para no decir nada del mundo real y no de problemas trasladados a una versión de juguete— más parecía que el éxito de aquel aprendiz darwiniano de Samuel se atribuyese a la relativa simplicidad del juego de damas y no al poder de la subyacente capacidad de aprendizaje. ¿Era éste, entonces, el final de la darwiniana IA? ¡Naturalmente que no! El programa de Samuel hibernó durante algún tiempo hasta que los ordenadores y científicos de la computación avanzaron hasta niveles de mayor complejidad.

Hoy, la progenie del programa de Samuel se está multiplicando tan rápidamente que al menos tres nuevas revistas han sido fundadas en los últimos dos años para disponer de foros de discusión: *Evolutionary Computation*, *Artificial Life* y *Adaptative Behaviour*. La primera de estas

revistas subraya las preocupaciones de la ingeniería tradicional: el uso de la evolución simulada como un método para extender el poder práctico de diseño de los programadores o ingenieros del *software*. Los «algoritmos genéticos» diseñados por John Holland (que trabajó con Art Samuel en la IBM en su programa de damas) han demostrado su potencia en el mundo «no sin sentido» del desarrollo del *software* y se ha mutado en un *filum* de variaciones de algoritmos. Las otras dos revistas se concentran en una investigación con más aroma biológico, en la que las simulaciones de los procesos evolutivos nos permiten realmente, por primera vez, *estudiar* el proceso del diseño biológico en sí mismo, *manipulándolo*; o mejor dicho, manipulando una simulación a gran escala del proceso evolutivo. Como ha escrito Holland, los programas de vida artificial nos *permiten* «rebobinar la cinta de la vida» y volverla a ver una y otra vez, en muchas variaciones.

6. *La hermenéutica de los artefactos o la ingeniería revertida*

La estrategia de interpretar los organismos como si fueran artefactos tiene mucho en común con la estrategia conocida por los ingenieros como *ingeniería revertida* (Dennett 1990b). Cuando la compañía Raytheon quiere construir un artefacto electrónico para competir con otro artefacto de la compañía General Electric, compra varios de los artefactos de la GE y procede a analizarlos: esto es la *ingeniería revertida*. Los hacen funcionar, los llevan al banco de prueba, los radiografían, los desarman y someten a cada una de sus partes a un análisis interpretativo. ¿Por qué utiliza la GE unos alambres tan gruesos? ¿Para qué sirven estos registros ROM extra? ¿Es ésta una doble capa de aislamiento y, si es así, cuál era la preocupación de los ingenieros? Hay que tener en cuenta que la hipótesis dominante es que todas estas preguntas, expresadas como «por qué», tienen su respuesta. Todo tiene su *raison d'être*; la GE no hace nada en vano.

Naturalmente, si la sabiduría de los ingenieros que practican esta reversión del proceso de construcción incluye una sana ayuda para sus propios conocimientos, reconocerán que la hipótesis que sostiene el carácter óptimo de todo lo que hace la GE es demasiado fuerte; a veces los ingenieros ponen cosas estúpidas e inútiles en sus diseños, a veces olvidan quitar cosas que ya no cumplen una función, en ocasiones pasan por alto atajos obvios retrospectivamente. A pesar de todo, el carácter óptimo del trabajo analizado debe ser la hipótesis que se mantenga; si los ingenieros que realizan esta

operación no pueden asumir que hay buenas razones para los hechos que observan, no pueden iniciar su análisis[78].

La revolución de Darwin no descarta la idea de la ingeniería revertida, pero, más que nada, permite que ésta sea reformulada. En lugar de tratar de deducir lo que Dios ha intentado, tratamos de deducir cuál es la razón, si hay alguna, por la que la «madre naturaleza» —es decir, el propio proceso de evolución mediante selección natural— «discierne» o «discrimina» hacer las cosas de una forma y no de otra. Algunos biólogos y filósofos se sienten molestos en conversaciones de este tipo acerca de las razones de la madre naturaleza. Piensan que éste es un paso atrás, una concesión sin motivo a los hábitos predarwinianos del pensamiento o, cuando menos, una traicionera metáfora. De este modo, tienden a coincidir con Tom Bethell, el reciente crítico del darwinismo, cuando piensa que hay algo improbable en este doble estándar (véase Bethell 1976:111). Yo sostengo que esta posición respecto a las relaciones entre evolución e ingeniería revertida no sólo está bien motivada, sino que es extremadamente fructífera y, de hecho, inevitable. Como ya hemos visto, incluso a nivel molecular, no se puede hacer biología sin hacer ingeniería revertida y no se puede hacer ingeniería revertida sin preguntarse por las razones que hay para cualquier cosa que estudiemos. Tenemos que preguntar mediante «por qués». Darwin no nos dijo que no lo hiciéramos, sino que nos enseñó cómo contestar (Kitcher 1985a).

Comoquiera que el próximo capítulo lo vamos a dedicar a la defensa de esta tesis, demostrando las circunstancias en las que el proceso de evolución por selección natural se comporta como un ingeniero *inteligente*, es importante que, en primer lugar, consideremos dos situaciones en las que el proceso evolutivo *no* actúa realmente como un ingeniero *inteligente*.

Cuando los seres humanos diseñamos una nueva máquina, empezamos generalmente teniendo a mano una buena versión de la máquina, sea un modelo más antiguo, una maqueta o un modelo a escala de lo que vamos a construir. Lo examinamos con cuidado y probamos varias modificaciones. «Si dobláramos un poco esta mordaza hacia arriba y desplazamos un poco este fragmento de cremallera, funcionaría aún mejor». Pero no es así como trabaja la evolución. Esto se ve con suma claridad a nivel molecular. Una molécula concreta tiene la *forma* que tiene, por lo que no tolera que se la doble ni que se la cambie de forma. Lo que la evolución debe hacer cuando mejora el diseño molecular es construir otra molécula, una que sea casi como la que no funciona bien y descartar simplemente la vieja.

Aunque se aconseja no fustigar *nunca* a los caballos a mitad de carrera, la evolución *siempre* los fustiga. La evolución no puede establecer algo si no es mediante selección o descarte. De este modo, en cada proceso evolutivo —y, en consecuencia, en toda verdadera explicación de la evolución— hay siempre un débil aunque desconcertante tufo a juego de azar. A este fenómeno lo llamaré *palo y zanahoria*, según la práctica deshonesta de atraer a clientes anunciando algún producto a precio de saldo y entonces, cuando ya han entrado a la tienda, se les trata de vender algo más caro. Al contrario de esta práctica, el método del *palo y la zanahoria* en la evolución no es realmente inicuo; aunque parece serlo, ya que no explica lo que en un principio querríamos. Cuando se le pregunta, cambia de tema hábilmente.

La más ominosa sombra del método del *palo y la zanahoria* la hemos analizado, en su forma más pura, en el capítulo 2, en la rara apuesta que podemos proponer a alguien: la de ganar diez lanzamientos consecutivos de monedas a cara o cruz sin perder una sola vez. No sabemos por adelantado quién va a ser el ganador: sólo sabemos que la corona de ganador pasará —ha de pasar como una necesidad del proceso algorítmico— a *uno* u *otro* tan pronto como yo ejecute el algoritmo. Si pasa por alto esta posibilidad y acepta mi atractiva apuesta es porque, en la práctica humana, estamos demasiados acostumbrados al seguimiento de individuos y de proyectos de construcción alrededor de individuos identificados y de sus posibilidades futuras. Y el ganador de la partida está equivocado si piensa que debe haber una explicación de por qué fue él quien ganó. No hay una razón para que gane él: sólo una muy buena razón para que gane alguien. Pero, siendo humano, el ganador no dudará en pensar que debía haber una razón por la que él ganó: «¡Si vuestro relato del proceso de la evolución no puede explicar esto, entonces está olvidando algo importante!». A lo que el evolucionista debe responder con calma: «Señor, sé qué es lo que usted está esperando, pero permítame tratar de interesarle en algo un poco más proporcionado, menos presuntuoso y un poco más defendible».

¿Se le ha ocurrido pensar la suerte que ha tenido por el hecho de vivir? ¡Más del 99 por ciento de todas las criaturas que han vivido han muerto sin progenie, pero ni uno sólo de nuestros antepasados se encuentra en este grupo! ¡Qué linaje real de ganadores ha llegado hasta cada uno de nosotros! (Naturalmente esto mismo ocurre con cualquier percebe, hoja de hierba o mosca). Algo que es aún más misterioso. ¿No es cierto que hemos aprendido que la evolución funciona eliminando lo inapropiado? Gracias a sus defectos de diseño estos perdedores tienen una «patética pero alabable tendencia a

morir antes de reproducirse en una progenie» (Quine 1969:126). Este es el verdadero motor de la evolución darwiniana. No obstante, si miramos atrás, a nuestro árbol familiar, con una visión de túnel, encontraremos una multitud de diferentes organismos, con una amplia variedad de puntos fuertes y puntos débiles pero, curiosamente, sus debilidades nunca conducen a uno sólo de ellos a una desaparición prematura. Desde esta perspectiva, parece que la evolución no pudiera explicar una característica singular que hubiéramos heredado de nuestros antecesores. Supongamos que miramos atrás, hacia el abanico de nuestros antepasados. Primero observamos que, eventualmente, la apertura en abanico se detiene y se dobla; cada uno de nosotros compartimos múltiples antepasados con cualquiera que esté vivo hoy, y hay muchos relacionados con muchos de nuestros antepasados. Cuando observamos la totalidad del árbol a lo largo del tiempo vemos que los últimos y más recientes antepasados poseen mejoras que les faltaban a los más antiguos, aunque todos los hechos cruciales —todos los acontecimientos de la selección — suceden fuera de la escena, entre bastidores: ninguno de nuestros antepasados, hasta llegar a las bacterias, sucumbió a la predación antes de reproducirse o bien quedó fuera de la competición para cruzarse.

La evolución explica todas las características que heredamos de nuestros antepasados, pero no explica por qué somos tan afortunados al tenerlas. Explica por qué los ganadores de hoy poseen las características que tienen, pero no por qué *estos individuos* poseen las características que tienen[79]. Consideremos ahora la siguiente situación: encargamos un nuevo coche y especificamos que lo queremos de color verde. En el día convenido vamos a recogerlo y allí está el coche verde y nuevo. ¿Cuál es la pregunta correcta que debemos hacer? «¿Por qué es verde este coche?» o «¿por qué está aquí este coche verde?». (En siguientes capítulos volveremos a ocuparnos de las implicaciones del método del *palo y la zanahoria*).

La segunda diferencia importante entre los procesos —y en consecuencia entre los productos— de la selección natural y de la ingeniería humana se relaciona con una característica de la selección natural que a mucha gente le parece muy paradójica: su absoluta falta de previsión. Cuando los ingenieros diseñan algo (estos ingenieros, en contraposición a la *ingeniería revertida*, hacen ingeniería hacia delante) deben de tener cuidado con un problema notorio: los efectos colaterales no previstos. Cuando dos o más sistemas, bien diseñados aisladamente, se reúnen en un supersistema, a menudo se producen interacciones que no forman parte del diseño que se intentó, sino que pueden ser positivamente nocivas; la actividad de un sistema daña inadvertidamente

la actividad del otro. El único método práctico para prevenir estos efectos colaterales imprevistos, ya que por su misma naturaleza son impredecibles para aquellos cuya mirada está forzosamente restringida a uno de los subsistemas diseñados, es diseñar todos los subsistemas de modo que tengan fronteras relativamente impenetrables, que coincidan con los límites epistémicos de sus creadores. Es característico de los ingenieros que intenten aislar los subsistemas unos de otros con el objetivo de conseguir un diseño global en el que cada subsistema tenga una función singular y bien definida dentro de la totalidad.

¡El conjunto de supersistemas que posee una fundamental y abstracta arquitectura es, naturalmente, vasto e interesante, aunque no incluye muchos de los sistemas diseñados por la selección natural! El proceso de selección carece evidentemente de previsión. Dado que no tiene previsión, los efectos imprevistos e imprevisibles no son tenidos en cuenta como tales por el proceso evolutivo; al contrario de lo que ocurre en los ingenieros, sigue adelante a través de un proceso derrochador que crea un enorme número de diseños relativamente *no* aislados, la mayoría de los cuales son fallos sin esperanza que se deben a defectos colaterales autodestructivos, aunque algunos, por un azar ciego, se libran de ese hado ignominioso. Sin embargo, esta aparentemente ineficiente filosofía del diseño acarrea un enorme dividendo que es casi inalcanzable para los más eficientes procesos, realizados de arriba abajo por los ingenieros: gracias a no mantener una actitud sesgada con respecto a los efectos colaterales no examinados, pueden extraer ventajas de los raros casos en los que parece que pueden emerger efectos colaterales beneficiosos, por la llamada *serendipity*. Algunas veces, aparecen diseños en los que los sistemas interactúan para producir algo más de lo que parecía ser el objetivo. En particular, pero no exclusivamente, en tales sistemas se producen elementos que cumplen múltiples funciones.

Es evidente que los ingenieros no desconocen los elementos con funciones múltiples, pero su relativa rareza se pone de manifiesto por el placer que sentimos cuando encontramos uno nuevo. (Mi elemento favorito que cumple funciones múltiples es la impresora portátil Diconix. Esta pequeña impresora funciona con baterías recargables bastante grandes que deben ser almacenadas en algún lugar; este lugar es la propia platina y el rodillo de la impresora donde quedan perfectamente ajustadas). Una breve reflexión nos hace ver que tales ejemplos de funciones múltiples son epistémicamente accesibles a los ingenieros bajo varias circunstancias propicias, pero también podemos ver que, de una manera general, tales

soluciones de diseño deben ser excepciones en un trasfondo de aislamiento estricto de elementos funcionales. En biología nos encontramos con aislamientos anatómicos nítidos para determinadas funciones (el riñón es completamente distinto del corazón, los nervios y los vasos sanguíneos son conductos separados que atraviesan el cuerpo humano, etc.) y sin este aislamiento fácilmente discernible en biología, la ingeniería revertida sería, sin duda, humanamente imposible. Pero también observamos superposición de funciones y en las que, además, algunos de los efectos más sobresalientes, observables en la interacción de estos elementos, pueden no ser realmente funciones, sino simplemente productos de desecho de las múltiples funciones cumplidas**[80]**.

Hasta hace muy poco tiempo, los biólogos que deseaban practicar la ingeniería revertida tenían que concentrarse en la resolución de las características del diseño de «productos acabados» como son los organismos. De tales productos, los biólogos pueden recolectar cientos o miles, estudiar sus variaciones, dividirlos y manipularlos *ad lib*. Resulta bastante más difícil conseguir una adquisición epistémica sobre el proceso de desarrollo o de construcción mediante el cual un genotipo consigue ser «expresado» en un fenotipo completamente formado. Y los procesos de *diseño* que dan forma a los procesos de desarrollo que moldean los «productos acabados» son muy inaccesibles a las observaciones intrusivas y a las manipulaciones que la buena ciencia (o la buena ingeniería revertida) aplican. Cabe considerar un esquema del proceso histórico de la evolución y hacer que se desarrolle rápidamente hacia delante (como las imágenes fotográficas de plantas creciendo o de los cambios de tiempo meteorológico, que son siempre una forma elegante de visualizar los modelos), pero no podemos *rebobinar la cinta* y observar la evolución de las condiciones iniciales. Ahora, gracias a las simulaciones por ordenador, es posible *estudiar* las hipótesis acerca de los procesos de diseño que han estado siempre latentes en el corazón de la visión darwiniana. No es de extrañar que sean más complicados y en sí mismos más intrincadamente diseñados de lo que habíamos pensado.

Una vez que los procesos de investigación más desarrollo [I+D] y de construcción aparecen en escena, se puede ver que la miopía que ha confundido a menudo a los hermeneutas de los artefactos humanos tiene múltiples paralelismos en la biología. Cuando nos interesamos por la hermenéutica de los artefactos, tratando de descifrar el diseño de objetos no cubiertos por los arqueólogos, o tratando de recuperar una interpretación apropiada de los antiguos monumentos con los que hemos crecido, existe una

tendencia a pasar por alto la posibilidad de que algunas de las características que nos confunden no desempeñen, en último término, una función en el producto acabado, aunque hayan tenido un papel funcional crucial durante el proceso que ha creado el producto.

Las catedrales, por ejemplo, muestran muchos hechos estructurales curiosos que han dado lugar a fantasías funcionales y debates enconados entre los historiadores del arte. Las funciones de algunos de estos hechos estructurales son, sin duda, obvias. Las numerosas escaleras que, en forma de espiral, suben por el interior de los pilares y de los muros son vías útiles para que los guardianes del templo alcancen las partes más altas del edificio: el tejado y el espacio entre la bóveda y el tejado donde está oculta la maquinaria que hace descender las lámparas hasta el suelo y puedan sustituirse las velas. Pero muchas de estas escaleras estarían allí incluso si estas posteriores vías de acceso no hubieran sido previstas por los constructores; no era simplemente el mejor, sino el único método que los constructores tenían a mano para que los operarios y los materiales llegaran a donde fuera necesario durante la construcción. Otros pasadizos que no conducían a ninguna parte estaban allí, probablemente, para que llegara el aire fresco al interior de las paredes (Fitchen 1961). El mortero medieval tardaba mucho tiempo —años, en algunos casos— en cuajar y cuando cuajaba se contraía, así que era necesario tomar precauciones para minimizar el grosor de las paredes, de modo que la distorsión fuera la mínima cuando el edificio se asentaba. (De este modo, aquellos —pasadizos cumplían una función similar a las «aletas» para la disipación del calor en el espacio donde se encuentra el motor del automóvil, con la diferencia de que sus funciones acababan una vez que la obra de los edificios había madurado).

Cuando contemplamos una catedral, examinada simplemente como un producto acabado, bastantes detalles sin interés aparente se nos presentan como extraordinariamente enigmáticos en cuanto nos preguntamos sobre cómo pudo construirse. En este terreno abundan problemas similares a la clásica pregunta sobre la primacía del huevo o la gallina. Si construimos los arbotantes antes de construir la bóveda central, éstos pueden empujar a las paredes hacia dentro; si se construye en primer lugar la bóveda antes de que los arbotantes se instalen, ésta separará las paredes; si tratamos de construir ambas estructuras al mismo tiempo, parece probable que la fase de construcción de una debe ir acompasada con la fase de construcción de la otra. Éste es, con seguridad, un problema que tiene una solución —o probablemente muchas soluciones diferentes— aunque imaginarla y luego

buscar la evidencia que la confirme o la rechace es un reto. Una de las estrategias que se repiten ya la hemos visto utilizada en la hipótesis de los cristales de arcilla de Cairns-Smith: en aquella situación debe haber habido elementos que sirvieron de andamiaje y que han desaparecido, ya que funcionaron sólo durante el proceso de construcción. Dichas estructuras dejan, a menudo, señales de su antigua presencia. Orificios en las paredes, llamados *mechinales*, son las señales más obvias. Los extremos de pesados troncos llamados *parales* se fijaban temporalmente en las paredes para sostener los andamios; sus extremos se introducían en los mechinales.

Muchos de los elementos decorativos de la arquitectura gótica, como los elaborados modelos de *nervios* en las bóvedas, son en realidad componentes estructuralmente funcionales, aunque cumplen su función sólo durante la fase de construcción. Los nervios tienen que levantarse antes de que los espacios de la bóveda, delimitados por ellos, se rellenen. Los nervios aportan rigidez al delicado tejido que va «centrándose» en el esqueleto de la bóveda, el cual, de otro modo, tendería en ceder y deformarse bajo el peso, temporalmente no equilibrado, de las partes ya construidas de la bóveda. Utilizando los métodos y materiales medievales, existían graves limitaciones a la resistencia del andamiaje que podía ser construido y mantenido en posición con seguridad a gran altura. Estos límites dictaron muchos de los detalles «ornamentales» de la iglesia terminada. Otra forma de subrayar este punto: muchos productos fácilmente *concebibles* como terminados fueron sencillamente imposibles de ser erigidos, debido a las limitaciones del proceso de construcción; y muchos de los detalles aparentemente sin función de los edificios existentes son, de hecho, características que han facilitado el diseño y sin las que el producto acabado podía no existir. La invención de las grúas (grúas reales y no metafóricas) y los artefactos de ellas derivados abrieron regiones anteriormente inaccesibles en el espacio de las posibilidades arquitectónicas**[81]**.

La cuestión es simple, pero nos proyecta una larga sombra: cuando planteamos cuestiones que afectan a las funciones de algo —sea organismo o artefacto— debemos recordar que éste ha llegado a su forma actual o final mediante un proceso que tiene sus propios requerimientos y éstos son exactamente tan receptivos al análisis funcional como cualesquiera otras características del estado final del producto considerado. No suenan las campanas para marcar el final de lo que es el proceso de construcción y el comienzo del funcionamiento (véase Fodor 1987:103). Los requerimientos

para que un organismo funcione en cada etapa de su vida aplican rígidas limitaciones sobre sus posteriores características.

D'Arcy Thompson [1917] dejó dicho que todo se debe a que ocurrió así, y sus propias reflexiones sobre el proceso histórico del desarrollo le condujeron a la promulgación de sus «leyes de la forma» que, a menudo, son citadas como ejemplo de leyes biológicas irreductibles a leyes físicas. Es innegable la importancia de estas reconstrucciones de los procesos de desarrollo y la investigación de sus implicaciones, pero este tema se encuentra, a veces, mal ubicado en aquellas discusiones que intentan *contrastar* tales limitaciones al desarrollo con los análisis funcionales. Ningún análisis funcional bien fundamentado está completo hasta que ha confirmado (como muchos de esos puntos pueden ser confirmados en cualquier momento) que ha sido especificado un método de construcción. Si algunos biólogos han pasado por alto este requerimiento, están cometiendo el mismo error que los historiadores del arte que ignoran el proceso de construcción de sus monumentos. Lejos de ser considerados con excesiva mentalidad de ingeniero, estos biólogos no se han tomado las cuestiones de ingeniería con la suficiente seriedad.

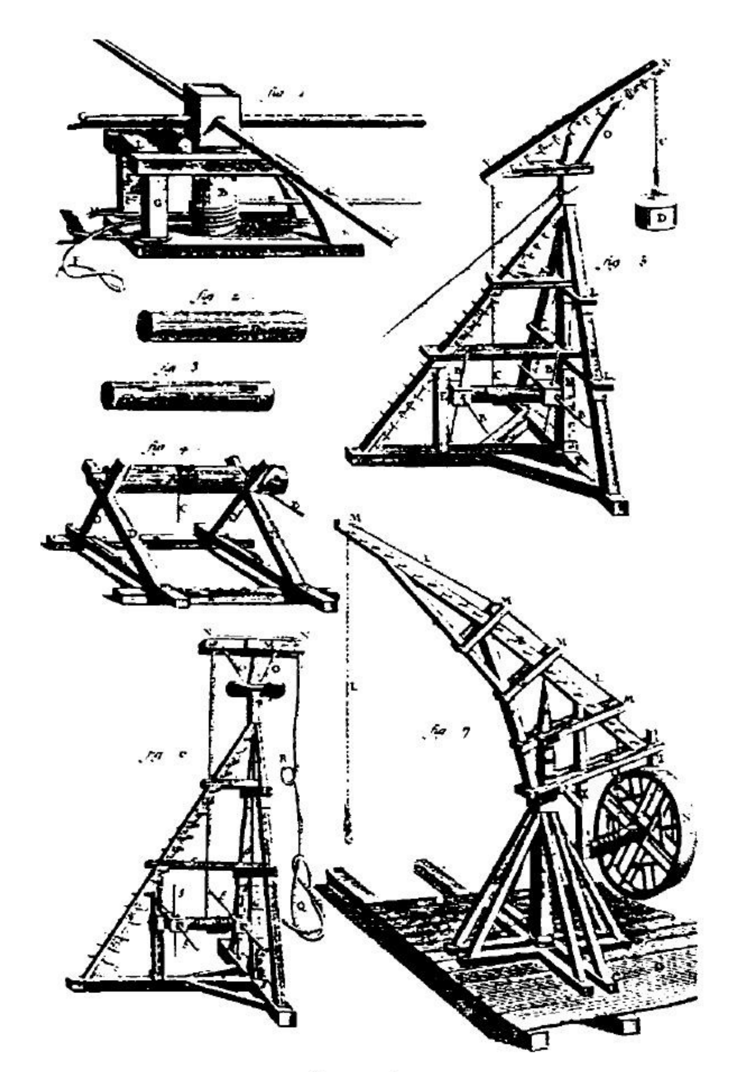


Figura 8.1
Primeras grúas giratorias y otros dispositivos para levantar o mover pesos. (De Diderot y D'Alembert, *Encyclopédie*, 1751-1772, reproducido en Fitchen 1986).

Figura 8.1

7. *Stuart Kauffman como metaingeniero*

A partir de Darwin hemos llegado a considerar a los organismos como artefactos montados chapuceramente y a la selección como la única fuente del orden. Con todo, Darwin no pudo haber empezado a sospechar del poder de la autoorganización. Debemos buscar nuestros principios de adaptación en nuevos sistemas complejos.

Stuart Kauffman
(Citado en Ruthen, «*Adapting to Complexity*»)

La historia tiende a repetirse. Hoy todos reconocemos que el redescubrimiento de las leyes de Mendel, y con ello el concepto de gen como una unidad de la herencia, fue la salvación del pensamiento darwiniano, aunque no como apareció en su momento. Como Maynard Smith ha subrayado, «el primer impacto del mendelismo sobre la biología evolucionista fue, sin duda, insignificante. Los primeros mendelianos se consideraban antidarwinianos» (1982:3). Esta fue una de las muchas supuestas revoluciones antidarwinianas que se han convertido en reformas prodarwinianas, sacando a la peligrosa idea de Darwin de uno u otro lecho de enfermo y recuperándola para el trabajo. Otra nueva dirección que se despliega hoy ante nuestros ojos es el pensamiento evolucionista encabezado por Stuart Kauffman y sus colegas del Instituto de Santa Fe. Como todo autobús de campaña electoral que se precie, tiene su eslogan: «Evolución al borde del caos». El nuevo libro de Kauffman, titulado *The Origins of Order: Self-Organization and Selection in Evolution* (1993), resume y describe ampliamente la investigación en la que se había comprometido desde hacía varias décadas y nos permite comprobar, por primera vez, cómo el propio Kauffman sitúa sus ideas en el contexto de la historia de este campo de ciencia.

Muchos lo califican como el asesino de Darwin, el que consiguió eliminar de la escena su opresiva presencia, haciéndolo, además, con la deslumbrante espada de una nueva rama de la ciencia: la teoría del caos y la teoría de la complejidad, *attractors* extraños y «fractales». El propio Kauffman había tenido esta tentación en el pasado (Lewin 1992:40-43), pero su libro se eriza con advertencias para mantener a raya el abrazo de los antidarwinianos. Kauffman comienza el prefacio de su libro describiéndolo como «un intento de incluir a Darwin en un contexto más amplio»:

> Por lo tanto, nuestra tarea no sólo es explorar las fuentes del orden que puede estar disponible para la evolución. También debemos integrar tal conocimiento en la visión básica ofrecida por Darwin. La selección natural, cualquiera que sean nuestras dudas en casos concretos, es seguramente una fuerza preeminente en la evolución. Por lo tanto, al combinar los temas de la autoorganización y de la selección, debemos extender la teoría evolucionista de tal modo que asiente sobre unos fundamentos más amplios y entonces levantar un nuevo edificio (Kauffman 1993:XIV).

He ido tan lejos citando a Kauffman en este punto porque también he sentido el vendaval del sentimiento antidarwinista entre mis propios lectores y críticos, y sé que ellos se sentirán muy motivados para sospechar que estoy simplemente reformando las ideas de Kauffman para que encajen en mi propio y sesgado punto de vista. ¡No!, el propio Kauffman —para cualquier cosa que valga la pena— ve ahora su trabajo como una profundización del

darwinismo, no como un rechazo. Pero, entonces, ¿cuál puede ser su punto de vista respecto a «la autoorganización espontánea» como «una fuente de orden» si no una tajante negación de que la selección sea la fuente fundamental del orden?

Ahora que es posible construir en los ordenadores escenarios de la evolución verdaderamente complejos, rebobinando la cinta una y otra vez, podemos ver modelos que eludían los primeros teóricos darwinistas. Lo que nosotros vemos, dice Kauffman, es que el orden «resplandece» *a pesar* de la selección, no a causa de ésta. En lugar de testificar el incremento gradual de la organización bajo la continua presión de la selección acumulativa, somos testigos de la incapacidad de la presión selectiva (la cual puede ser cuidadosamente manipulada y monitorizada en las simulaciones) para sobrepasar una inherente tendencia de la población en cuestión a transformarse a sí misma en modelos ordenados. Todo esto parece ser, a primera vista, una demostración brillante de que la selección natural no puede ser, después de todo, la fuente de la organización y el orden; esto sería, naturalmente, la ruina de la idea darwiniana.

Aunque, como hemos visto, hay otra manera de ver la cuestión. En efecto, ¿qué condiciones deben existir para que ocurra la evolución por selección natural? Las palabras que puse en boca de Darwin eran simples: dadme orden y tiempo y os daré diseño. Pero lo que hemos aprendido más tarde es que no toda variedad de orden es suficiente para que exista capacidad de evolución. Como hemos visto ilustrado por el Juego de la Vida de Conway, tenemos que tener el apropiado tipo de orden, con la mezcla apropiada de libertad y treno, crecimiento y decadencia, rigidez y fluidez, para que sucedan buenas cosas. Sólo alcanzaremos la evolución, como el eslogan del Instituto de Santa Fe proclama, en el borde del caos, en las regiones de posible ley que forma la híbrida zona entre el orden rígido y el caos destructivo. Afortunadamente, nuestra porción del universo está situada justamente en esta zona, en la que las condiciones para la evolución están correctamente ajustadas. ¿Y de dónde proceden estas saludables condiciones? En principio, éstas pueden proceder de la sabiduría y la previsión de un diseñador como Conway, o de un proceso evolutivo previo, ya se trate de un proceso con selección o sin ella. En efecto —y éste es, creo, el centro de la visión de Kauffman—, la capacidad de evolucionar en sí misma no sólo debe evolucionar (para nosotros significa estar aquí), sino que es *probable* que evolucione, es casi seguro que evolucione, debido a que se trata de un movimiento forzado en el juego del diseño[82]. O encontramos el camino que conduce a la capacidad de

evolucionar o no vamos a ninguna parte, aunque hallar ese camino no es, en sí mismo, un gran asunto; es algo «obvio». Los principios de diseño que hacen posible la evolución biológica los encontraremos siempre, una y otra vez, sin importar cuántas veces rebobinamos la cinta. «Contrariamente a todas nuestras expectaciones, la respuesta, pienso, es que puede ser sorprendentemente *fácil*» (Kauffman 1993:XVI).

Cuando en el capítulo 6 considerábamos las fuerzas que se mueven en el espacio de diseño estábamos pensando en las características de los productos finales que eran tan obviamente «correctos» que no nos sorprendería encontrarlos por separado: por ejemplo, la aritmética entre alienígenas inteligentes y los ojos dondequiera que exista locomoción a través de un medio transparente. Pero ¿qué hay con respecto a las características del proceso de creación de aquellos productos? Si hay reglas fundamentales acerca de cómo las cosas han de ser diseñadas y acerca del orden en el que las innovaciones del diseño pueden ser creadas, las estrategias del diseño que están ligadas a funcionar o a fracasar, entonces éstas deben ser dirigidas por la evolución con tanta seguridad como lo son las características de los productos finales. Asumo que lo que Kauffman ha descubierto no son *leyes de la forma* sino *reglas para el diseño*: los imperativos de la metaingeniería. Kauffman tiene muchas observaciones efectivas que hacer respecto a los principios de la metaingeniería que gobiernan el proceso mediante el cual pueden ser creados, en principio, nuevos diseños. Podemos considerar que son características del fenómeno global de la evolución que ha sido ya descubierto, que, en efecto, ha conseguido fijarse en nuestra parte del universo. (No nos sorprendería encontrarlos en cualquier parte del universo donde hay cosas diseñadas, porque es la única vía de diseñar cosas).

La evolución adaptan va es un proceso de búsqueda —impulsado por mutación, recombinación y selección— de *paisajes de adaptabilidad* al medio fijos o en formación. Bajo estas fuerzas, una población en adaptación fluye sobre el paisaje. La estructura de tales espacios, lisa o rugosa, gobierna tanto la capacidad de evolución de la población como la adaptabilidad sostenida de sus miembros. La estructura del *paisaje de adaptabilidad* impone, inevitablemente, limitaciones en el espacio adaptativo (Kauffman 1993:118).

Obsérvese que todo esto es puro darwinismo; todo ápice aceptable y no y revolucionario, pero con un mayor grado de énfasis sobre el papel de la topología del *paisaje adaptativo*, que, como argumenta Kauffman, tiene un profundo efecto sobre el *ritmo* según el cual pueden ser encontradas innovaciones del diseño y el *orden* según el cual los azares del diseño pueden acumularse. Si alguna vez hemos intentado escribir un soneto, nos hemos

enfrentado con el problema básico de diseño que los modelos de Kauffman examinan: la «epistasis», o sea, la interacción entre genes. Como el poeta novato pronto descubre, ¡no es fácil escribir un soneto! Decir algo con significado —no sólo bello— dentro de los rígidos requerimientos de la forma del soneto es un ejercicio frustrante. Tan pronto como terminamos el primer intento de una línea, tenemos que revisar muchas de las otras líneas y esto nos obliga a abandonar algunas de las excelencias conseguidas con dificultades y así sucesivamente, dando vueltas y más vueltas, buscando una forma global correcta o tratando de hallar la mejor forma posible. El matemático Stanislaw Ulam comprobó que las dificultades de la poesía pueden ser una fuente de creatividad y no un freno. La idea puede aplicarse, justamente por la misma razón, a la creatividad de la evolución:

Cuando era niño me di cuenta de que la función de la rima en la poesía consistía en obligarnos a encontrar lo no obvio a causa de la necesidad de hallar una palabra que rimase. Esta necesidad fuerza nuevas asociaciones y casi garantiza desviaciones de cadenas de rutina o de pensamientos encadenados. Paradójicamente se convierte en una suerte de mecanismo automático de originalidad (Ulam 1976:180).

Antes de Kauffman, los biólogos tendían a dejar de lado la perspectiva de que la evolución debería confrontar el mismo tipo de interacciones penetrantes, ya que no se disponía de un método claro para estudiarla. Su trabajo demostró que hacer viable un genoma es más parecido a escribir un buen poema que anotar simplemente la lista de la compra. Dado que la estructura de los *paisajes adaptativos* es más importante de lo que habíamos pensado (con nuestros paisajes más simples como los modelos Mount Fuji de escalada conocidos como *hill-climbing*) hay limitaciones o *métodos* para mejorar los diseños que mantienen a los proyectos de ingeniería canalizados en caminos más definidos de lo que habíamos imaginado, para conseguir resultados.

La capacidad de evolucionar, la capacidad de buscar una razonable fracción del espacio, puede ser optimizada cuando la estructura del paisaje, la tasa de mutación y el tamaño de la población están ajustados, de tal modo que las poblaciones comienzan justamente a «mezclarse» desde regiones locales de este espacio [Kauffman 1993:95].

Una característica ubicua en la evolución biológica fue resumida por Kauffman en el principio que dice que las «reglas generales generan el orden global». Éste no es un principio que gobierne la ingeniería humana. Las pirámides se construyen siempre desde la base, pero la organización del proceso constructivo, desde la época de los faraones, ha sido de arriba abajo, sometida a la dirección de un único autócrata que tenía una clara y, literalmente, dominante visión de la totalidad, pero que probablemente era

poco precisa en lo que se refiere a cómo debían realizarse los detalles locales. La dirección «global» ejercida desde arriba pone en movimiento una cascada jerárquica de proyectos «locales». Es una característica común de los proyectos humanos a gran escala que tengamos muchas dificultades para imaginar alternativas (Papert 1993, Dennett 1993a). Puesto que no reconocemos el principio descrito por Kauffman como familiar a la ingeniería, no estamos en condiciones de considerarlo como un principio de ésta, aunque yo sugiero que sí lo es. Ligeramente reformulado, podemos expresarlo así: hasta que tratamos de desarrollar organismos comunicantes que pueden formar grandes organizaciones de ingeniería, estamos ligados al siguiente principio preliminar del diseño: todo orden global debe ser generado por reglas locales. De este modo, todos los productos primarios del diseño, hasta la creación de algo dotado de algunos de los talentos organizadores del *Homo sapiens*, deben obedecer a algunas limitaciones nacidas de la «decisión de mando» de que todo orden debe ser conseguido mediante reglas locales. Todos los «intentos» para crear formas vivas que violen este precepto terminarán en un inmediato fracaso, o dicho con mayor precisión, no se iniciarán lo suficiente como para ser detectables como intentos.

Si no se recuerda, como he dicho, marcar el momento en que termina el proceso de investigación más desarrollo [I+D] y comienza la vida del «producto acabado», puede ser difícil de establecer, al menos algunas veces, si un determinado principio de diseño es un principio de ingeniería o un principio de metaingeniería. Un caso pertinente es la propuesta de Kauffman [1993] de llevar a cabo una rederivación de «las leyes de Von Baer» de la embriología. Una de las características más llamativas en los embriones de los animales es el hecho de que todos empiezan con un patrón bastante similar.

En su inicio, los embriones de peces, ranas, pollos y seres humanos son llamativamente similares… La explicación habitual, según las leyes de Von Baer, es que los mutantes (creo que él quiere decir «mutaciones») que afectan a la ontogenia precoz son más destructivos que las mutaciones que afectan a la ontogenia tardía. De este modo, es menos probable que se acumulen las mutaciones que alteran el desarrollo embrionario inicial, y los embriones en su inicio permanecen más similares entre uno y otro orden de organismos que los embriones más avanzados en su desarrollo. ¿Es este plausible argumento realmente tan plausible? (Kauffman 1993:75).

El darwiniano tradicional, según la lectura de Kauffman, coloca la responsabilidad de las leyes de Von Baer en un «mecanismo especial» constituido dentro de los organismos. ¿Por qué no observamos muchos productos acabados con embriones en estadios precoces marcadamente diferentes? Pues bien, dado que las órdenes de cambios que afectan a las fases precoces del proceso tienden a ser de efectos más destructivos sobre el

producto acabado que las órdenes de cambios que afectan a fases tardías del proceso, la madre naturaleza ha diseñado un mecanismo especial de desarrollo para protegerse contra tal experimentación. (Esto sería análogo a que la IBM prohibiera a sus científicos de la computación investigar arquitecturas alternativas a su chip de la unidad central de procesamiento (CPU); es decir, el *diseño* de una resistencia al cambio).

Y, por contraste, ¿cuál es la explicación de Kauffman? Empieza con el mismo punto pero toma una dirección diferente:

> Un bloqueo del desarrollo inicial, y en consecuencia de las leyes de Von Baer, no representa un mecanismo especial de canalización del desarrollo, cuyo sentido habitual es el bloqueo del fenotipo ante la alteración genética… En su lugar, el bloqueo del desarrollo inicial es una reflexión directa del hecho de que el número de caminos a seguir para mejorar los organismos alterando su ontogenia inicial ha disminuido más rápidamente que el número de caminos para mejorar alterando el desarrollo tardío (Kauffman 1993). (Véase también Wimsatt 1986:77).

Reflexionemos brevemente sobre este tema desde el punto de vista de la ingeniería humana. ¿Cuál es la razón de que los cimientos de las iglesias sean más parecidos que las plantas superiores? El darwinismo tradicional dice que en primer lugar deben construirse los cimientos y cualquier constructor inteligente nos diría que si *debemos* trabajar con elementos de diseño, trabajemos en primer lugar en los ornamentos de los capiteles o las ventanas, ya que existen menos posibilidades de que se produzca un hundimiento desastroso, que si tratamos de aplicar un nuevo método en la preparación de los cimientos. De este modo no es extraño que todas las iglesias en sus inicios se parezcan mucho, apareciendo las grandes diferencias en las fases más tardías del proceso de la construcción. En realidad, dice Kauffman, no hay muchas soluciones diferentes posibles al problema de los cimientos, como las hay para los posteriores problemas de la construcción. Incluso los constructores estúpidos que no aceptaron por cabezonería este hecho durante mucho tiempo, no llegaron a conseguir una amplia variedad de diseños para los cimientos. Estas diferencias de énfasis pueden parecer pequeñas aunque tienen importantes implicaciones. Kauffman dice que no es necesario buscar un mecanismo de canalización para explicar este hecho; se cuidará a sí mismo. Pero hay también un acuerdo subyacente entre Kauffman y la tradición que él desea suplantar: sólo hay muchas buenas vías para construir cosas, dadas las limitaciones del comienzo, y la evolución las encuentra una y otra vez.

La *no opcionalidad* de estas «elecciones» es la que Kauffman desea subrayar, y, así, él y su colega Brian Goodwin (por ejemplo 1986) están

particularmente impacientes por desacreditar la poderosa imagen, en un principio popularizada por los grandes biólogos franceses Jacques Monod y François Jacob, de la madre naturaleza como un artesano, embarcado en ese tipo de construcción de artilugios caseros que los franceses denominan *bricolage*. El primero que puso de moda el término fue el antropólogo Claude Lévi-Strauss (1966). Un *bricoleur* es un oportunista constructor de artilugios, un *satisficer* (Simon 1957) siempre dispuesto a apostar por la mediocridad si es lo suficientemente barata. Un artesano no es un profundo pensador. Los dos elementos clásicos del darwinismo que Monod y Jacob condensaron son el azar por un lado y por otro la absoluta falta de dirección y la miopía (o ceguera) del relojero por otro. Pero, dice Kauffman, «la evolución no es precisamente que *el azar consiga el ala*. No es el montaje de algo *ad hoc*, no es *bricolage*, no es un artefacto. Es el orden emergente honrado y pulido por la selección» (Kauffman 1993:644).

¿Está diciendo Kauffman que el relojero no es ciego? Naturalmente que no. Pero ¿qué dice entonces? Dice que hay principios de orden que gobiernan el proceso de diseño y que fuerzan la mano del artesano. Magnífico. Incluso un artesano ciego encontrará los movimientos forzados; éste no es el caso de un científico espacial, como suele decirse. Un artesano que no puede encontrar los movimientos forzados no merece ser un maldito artesano y no desea diseñar una cosa. Kauffman y sus colegas han hecho una importante serie de descubrimientos, pero el ataque a la imagen del artesano está, en gran parte, yo diría que fuera de lugar. El artesano, dice Lévi-Strauss, está deseando ser guiado por la naturaleza del material, mientras que el ingeniero desea que el material sea perfectamente maleable, como el hormigón tan apreciado por los arquitectos de la Bauhaus. Así el artesano es un profundo pensador después de todo, que se amolda a las limitaciones, sin luchar contra ellas. El ingeniero verdaderamente sabio no trabaja *contra naturam sino secundum naturam*.

Una de las virtudes del ataque de Kauffman es que llama la atención sobre una posibilidad menospreciada, una que podemos convertir en vívida con la ayuda de un ejemplo imaginario extraído de la ingeniería humana. Supongamos que la Acme Hammer Company descubre que los nuevos martillos fabricados por su rival, la Bulldog Hammer Inc., tienen mangos de plástico con los mismos intrincados modelos de espirales de colores que luce el nuevo modelo Acme Zeta. ¡Ladrones!, gritaron los representantes legales. ¡Habéis copiado nuestros diseños! Puede ser, pero también puede que no sea así. Puede ser, justamente, que sólo haya un método para hacer mangos de

plástico dotados de alguna resistencia y que éste consista en mezclar el plástico de esta forma. El resultado es inevitablemente un modelo con espirales. Sería casi imposible fabricar el mango de plástico de un martillo que pudiera ser utilizado sin que tuviera esas espirales y el descubrimiento de este hecho podría ser uno de los que se imponen en cualquiera que trate de hacer un mango de plástico para un martillo. Esto explicaría, por otro lado, la sospechosa similitud sin ninguna hipótesis de «descenso» o copia. Ahora bien, puede ser que la gente de la Bulldog haya copiado el diseño de Acme, y éstos *lo habrían encontrado en cualquier caso*, más pronto o más tarde. Kauffman puntualiza que los biólogos tienden a pasar por alto este tipo de posibilidad cuando deducen sus inferencias acerca del descenso y Kauffman llama la atención sobre muchos casos apremiantes en el mundo de la biología en los que la similitud de patrones no tiene nada que ver con el descenso. (Los casos más llamativos que Kauffman discute están ilustrados con los trabajos de Turing de 1952 sobre el análisis matemático de la creación de patrones espaciales en la morfogénesis).

En un mundo con principios de diseño no descubribles, todas las similitudes son sospechosas de ser copias (plagio o descenso).

Tenemos que llegar a pensar que la selección es, esencialmente, la única fuente de orden en el mundo biológico. Si «sólo» es una aseveración exagerada, entonces seguramente es preciso establecer que la selección se ve como la fuente dominante del orden en el mundo biológico. De esto se deduce que, desde el punto de vista actual, los organismos son la solución *ad hoc* para problemas de diseño conjuntados por selección. De esto se sigue también que la mayoría de las propiedades que están extendidas en los organismos lo están en virtud de un descenso común de un ancestro creado artesanalmente con mantenimiento selectivo de los artefactos útiles. De esto se deduce que nosotros vemos a los organismos como arrolladores accidentes históricos contingentes, inducidos por diseño (Kauffman 1993:26).

Kauffman desea subrayar que el mundo biológico es más un mundo de descubrimientos newtonianos (tal como el mundo de Turing) que un mundo de creaciones shakespearianas y ciertamente ha encontrado algunas excelentes demostraciones que apoyan esta aseveración. Pero me temo que su ataque a la metáfora del artesano alimente los ardientes deseos de aquellos que no aprecian la peligrosa idea de Darwin; que les traiga la falsa esperanza de que en los trabajos de la naturaleza están viendo no la forzada mano del artesano, sino la divina mano de Dios.

El propio Kauffman ha denominado a lo que él está haciendo «una indagación sobre la física de la biología» (Lewin 1992:43) y esto no está realmente en conflicto con lo que yo denomino metaingeniería. Es la investigación de las limitaciones más generales sobre los procesos que pueden

conducir a la creación y reproducción de las cosas diseñadas. Pero cuando Kauffman declara que esto es una búsqueda de «leyes», está alimentando un prejuicio antiingeniería (o podríamos decir la envidia frente a la física) que tanto distorsiona el pensamiento filosófico acerca de la biología.

¿Supone alguien que hay *leyes* de la nutrición? ¿Supone alguien que hay *leyes* de la locomoción? Hay toda clase de condiciones sobre la nutrición y la locomoción de límites muy imperturbables, y hay muchas regulaciones, reglas, trueques y cosas similares que se encuentran en cualquier mecanismo nutritivo o locomotor. Pero éstas no son leyes. Son como las fuertes regulaciones de la ingeniería de los automóviles. Consideremos la regularidad o constancia de que (*ceteris paribus*) la ignición esté asociada sólo con el uso de una llave. Existe una razón para esta regularidad, naturalmente, y ésta tiene que ver con el valor de los automóviles, su atracción para los ladrones, las opciones coste-eficacia (aunque no infalibles) facilitadas por la preexistente tecnología de la cerraduras, y así sucesivamente. Cuando uno comprende los miles de análisis de coste-beneficio realizados previamente a las decisiones sobre el diseño que se producen en la fabricación de automóviles se aprecia este tipo de constancias. No es ningún tipo de ley; es una regularidad que tiende a establecer un grupo complejo de competitivos *desiderata* (también conocidos como normas). Estas generalizaciones altamente fiables, que trazan las normas, no son leyes de ingeniería de automóviles, no son las contrapartidas biológicas de las leyes de la nutrición y locomoción. La localización de la boca en el extremo anterior y no en el extremo posterior de los organismos locomotores (*ceteris paribus*, ¡hay excepciones!) es una profunda regularidad, pero ¿por qué llamarla ley? Comprendemos *por qué* esto debe ser así, porque vemos que las bocas —o cerraduras o llaves— son *para* algo y porque ciertas soluciones son las mejores para conseguir un determinado objetivo.

CAPÍTULO 9
La búsqueda de la calidad

I. *El poder del pensamiento adaptacionista*

> «Desnudo como lo quiere la Naturaleza» era el persuasivo eslogan del movimiento naturista en sus comienzos. Pero la intención original de la Naturaleza era que la piel de los primates no estuviera desnuda.
>
> Elaine Morgan, *The Scars of Evolution*

> Juzgar un poema es como juzgar un pudin o una máquina. Se le pide que funcione. Sólo cuando un artefacto funciona, se infiere la intención de un artífice.
>
> W. Wimsatt y M. Beardsley,
> *«The Intentional Fallacy»*

Si conocemos de modo aproximado el diseño de un artefacto, podemos predecir su conducta sin preocuparnos de la física que subyace en sus componentes. Incluso los niños pequeños pueden fácilmente aprender a manipular objetos tan complicados como un vídeo sin conocer el principio según el cual funciona; saben lo que sucede cuando aprietan una secuencia de botones, porque saben lo que está previsto en su diseño. Los niños actúan a partir de lo que llamo *la posición para entender el diseño*. El técnico que repara los vídeos conoce muchas cosas más acerca del diseño del VCR («Video Casette Recorder») y sabe, hasta cierto punto, cómo interactúan sus componentes para producir tanto una función apropiada como una función anormal o patológica, pero también puede encontrarse completamente al margen del conocimiento de la física de estos procesos. Tan sólo los diseñadores de los VCR conocen sus principios físicos; ellos son los que deben descender a la que denomino la *posición para entender la física* con el fin de deducir qué tipos de revisiones del diseño pueden incrementar la calidad de la imagen, disminuir el deterioro por el uso y las rupturas de la cinta, o reducir el consumo de electricidad del equipo. Pero cuando estos diseñadores se implican en la ingeniería revertida —por ejemplo, los

productos de algún otro fabricante de vídeos— se aprovechan no sólo de la *posición para entender la física* sino también de lo que califico como *posición intencional* si tratan de deducir *lo que los diseñadores tenían en mente* cuando hicieron su trabajo. Entonces tratan al artefacto sometido a examen como el producto del proceso de desarrollo *razonado* de un diseño, a través de una serie de *elecciones* entre diversas alternativas, en las cuales las *decisiones* tomadas son aquellas que los diseñadores *consideraron mejores*. Reflexionar sobre las funciones postuladas para los distintos componentes del artefacto es establecer hipótesis acerca de las *razones* de la presencia de dichos componentes, lo que permite, a menudo, dar saltos gigantescos mediante deducciones que eluden la ignorancia que se tiene de la física subyacente o acerca de elementos del diseño del objeto del más bajo nivel.

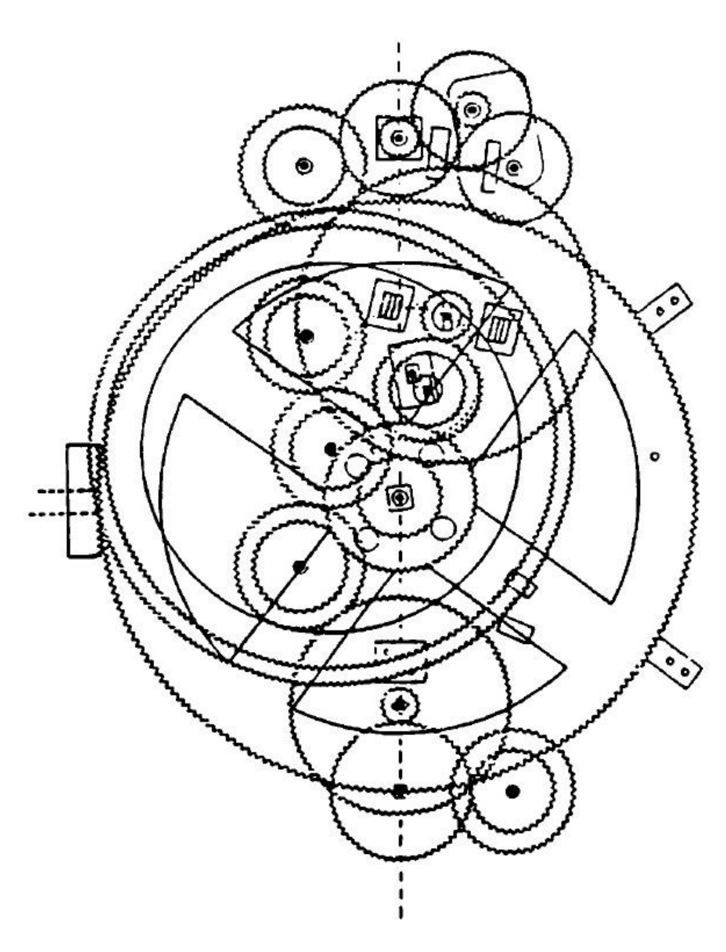


Figura 9.1 - Diagrama que representa el sistema de ruedas del mecanismo del Antikythera según Derek de Solla Price (Universidad de Yale).

Los arqueólogos y los historiadores encuentran a veces artefactos cuyo significado —función o propósito— resulta especialmente oscuro. Resultará instructivo examinar brevemente unos cuantos ejemplos de estas

interpretaciones o hermenéuticas de artefactos, para entender cómo se razona en tales casos[83].

El mecanismo del Antikythera descubierto en 1900 entre los restos de un naufragio que data de la antigua Grecia, es un montaje de ruedas dentadas de bronce, impresionantemente complejo. ¿Para qué servía? ¿Era un reloj? ¿Era la maquinaria para mover una estatua autómata, como las maravillas de Vaucanson del siglo XVIII? Era, casi con certeza, un planetario o mecanismo de relojería concebido para mostrar los movimientos de los planetas, y la prueba es que sería un *buen* planetario. Es decir, los cálculos de los períodos de rotación de sus ruedas conducen a una interpretación que habría sido una precisa representación (ptolomeica) de lo que entonces se conocía acerca del movimiento de los planetas.

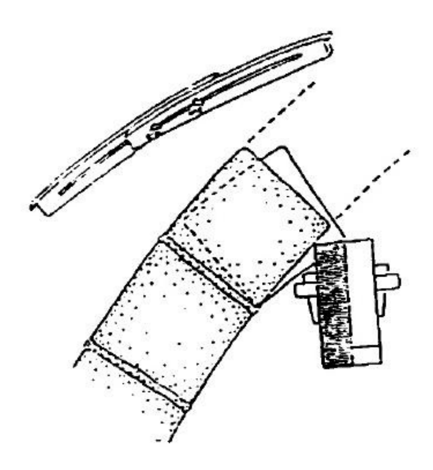


Figura 9.2

El dispositivo llamado *cerce* [calibrador] interpretado por Viollet-le-Duc como un soporte para las mampuestas: partes de la bóveda que se iban construyendo entre sus nervios. Arriba, el dibujo a menor escala muestra un *cerce* basado en la representación y descripción de Viollet-le-Duc. En la posición extendida del dispositivo se observa claramente cómo una tablilla con su ranura se solapa sobre la otra. Abajo (en la sección del dibujo con más detalles), con el *cerce* suspendido verticalmente como soporte de los bloques de piedra, se observa que, en estas circunstancias, los bloques de piedra de la bóveda no pueden mantener la alineación: aquellos que se apoyan contra la tablilla lejana (dibujada solo en su contorno) se inclinan más que aquellos que se apoyan contra la tablilla cercana (dibujada con sombreado). Como quiera que no se produjeron tales pérdidas de alineación en las mampuestas construidas, es obvio que este dispositivo no fue utilizado con este fin, a pesar de la afirmación de Viollet-le-Duc (Fichren 1961:101).

El gran historiador de la arquitectura Viollet-le-Duc describió un objeto denominado *cerce* —un calibrador— utilizado de algún modo en la construcción de las bóvedas de las catedrales. Viollet-le-Duc estableció la

hipótesis de que se trataba de una pieza movible de andamiaje utilizada como soporte provisional para las partes de la bóveda que se iban construyendo entre sus nervaduras, pero un investigador posterior, John Fichten (1961), argumentó que ésta no podía haber sido su función. Porque el *calibrador* no habría sido lo bastante fuerte en su posición extendida y, como muestra la figura 9.2, su utilización hubiese creado irregularidades, que no se encuentran en los espacios entre las nervaduras de las bóvedas de las catedrales. El extenso y elaborado argumento de Fitchen concluye que el *cerce* de Viollet-le-Duc no era más que una plantilla ajustable, conclusión que sostiene aportando una solución más elegante y versátil al problema del soporte provisional de las mampuestas construidas entre las nervaduras de la bóveda.

La característica más atractiva de estos argumentos es la confianza en las consideraciones que se hacen acerca de la optimización; esto ha de tenerse en cuenta contra la hipótesis que sostiene, por ejemplo, que algo es una deshuesadora de cerezas si pudo haber sido una deshuesadora de cerezas manifiestamente inferior. En ocasiones, un artefacto pierde su función original para asumir una nueva. La gente compra planchas de hierro colado pasadas de moda no para planchar la ropa, sino para usarlas como soportes o apoyos de libros o puertas; un elegante bote de mermelada puede convertirse en un recipiente para lápices y una trampa para langostas puede convertirse en un tiesto para plantas colocadas en el exterior. El hecho es que las planchas de hierro funcionan bastante mejor como soportes de libros que como planchas de ropa, si se las compara con la competencia a la que se enfrentan hoy. El procesador de un ordenador Dec-10 funciona hoy excelentemente a guisa de pesada ancla o «muerto» para mantener en posición la hoya de una barca. Ningún artefacto es inmune a tales apropiaciones y, no obstante, su propósito *original* puede «leerse» en su forma actual, su nuevo fin puede relacionarse con el original por un mero accidente histórico: el propietario del procesador obsoleto necesitaba con urgencia un ancla para la boya y, de manera oportunista, se sirvió del viejo procesador para este fin.

Las claves de estos procesos históricos serían simplemente ilegibles sin la hipótesis acerca de la optimización del diseño. Consideremos el llamado procesador de textos, es decir, la barata, portátil y glorificada máquina de escribir que utilizaba un disco de almacenaje y una pantalla electrónica, pero que no puede ser utilizado como ordenador en toda regla. Si abrimos uno de estos artefactos comprobaremos que están gobernados por una unidad central de procesamiento (CPU), como, por ejemplo, un chip 8088; se trata de un ordenador potente, muchísimo más potente, suave y versátil que el mayor

ordenador que Alan Turing hubiera visto jamás, aunque limitado a un servicio mínimo, a realizar una minúscula fracción de las tareas para las que *podía* haber sido aprovechado. ¿A qué se debe este exceso de capacidad funcional? Imaginemos a unos ingenieros marcianos realizando ingeniería revertida en este caso; quedarían desconcertados, aunque hay, naturalmente, una sencilla explicación histórica: la genealogía del desarrollo de los ordenadores disminuyó gradualmente los costes de manufactura de los *chips*, a tal punto que era más barato instalar un computador completo como *chip* en este procesador de textos, que construir un controlador de su circuito especialmente para este propósito. Téngase en cuenta que la explicación es histórica pero también procede ineludiblemente de una posición intencional. Diseñar procesadores de textos de esta forma fue una sabia decisión, ya que el análisis coste-beneficio demostró que esta opción era la *mejor* y la *más barata para resolver el problema*.

Es sorprendente lo poderosa que puede ser la posición intencional cuando se realiza ingeniería revertida, no sólo en artefactos humanos sino también en organismos. En el capítulo 6 analizamos el papel del razonamiento práctico —en especial el del análisis coste-beneficio— en la distinción de los movimientos forzados de lo que podríamos llamar los movimientos *ad lib*, y comprobamos cómo podía ser predecible que la madre naturaleza descubriese, una y otra vez, los movimientos forzados. La idea de imputar las «razones que estaban en el aire» al proceso no inteligente de la selección natural causa vértigo, aunque no significa negar los frutos de la estrategia. En los capítulos 7 y 8 hemos visto cómo la perspectiva de la ingeniería inspira la investigación en cada nivel biológico, a partir de las moléculas, y cómo esta perspectiva *siempre* implica distinguir lo mejor de lo peor y las razones que la madre naturaleza ha encontrado para esa distinción. La posición intencional es, por lo tanto, la palanca crucial en todos los intentos de reconstruir el pasado biológico. ¿Remontó el vuelo alguna vez el *Archaeopteryx*, la extinta criatura parecida a un pájaro que algunos han llamado dinosaurio alado? Nada más efímero, y menos indicado para dejar una huella fósil, que un vuelo a través del aire, pero si hacemos un análisis de ingeniería de sus garras veremos que están excelentemente adaptadas para agarrarse a las ramas y no para correr. Un análisis de la curvatura de las garras, complementado con el estudio aerodinámico de la estructura del ala del *Archaeopteryx* demuestra claramente que esta criatura estaba *bien diseñada* para volar (Feduccia 1993). De modo que, casi con certeza, voló o tuvo ancestros que volaron (no debemos olvidar la posibilidad de que persistiera una funcionalidad excesiva

como es el caso del ordenador en el procesador de textos). La hipótesis de que el *Archaeopteryx* voló no ha sido todavía totalmente confirmada, y no convence a todos los expertos, pero sugiere muchas preguntas dirigidas al fósil, y cuando se siga la pista de esas preguntas la evidencia aumentará a favor o en contra de la hipótesis. La hipótesis puede someterse a prueba.

La palanca de la ingeniería revertida no es justamente para poner de manifiesto los secretos de la historia; aún es más espectacular profetizando secretos inimaginables del presente. ¿Por qué hay colores? La codificación mediante colores es generalmente interpretada como una reciente innovación de la ingeniería, pero no es así. La madre naturaleza la descubrió muchísimo antes (detalles sobre esta cuestión se encuentran en la sección sobre *por qué hay colores* en Dennett 1991a:375-383). Nuestros conocimientos sobre esta cuestión se deben a las líneas de investigación abiertas por Karl von Frisch y, como Richard Dawkins ha señalado, fue Von Frisch quien utilizó un audaz ejercicio de ingeniería revertida para dar el primer paso.

Como desafío a Von Hess, con su prestigiosa ortodoxia, Von Frisch (1967) demostró, de manera concluyente, mediante experimentos controlados, que los peces y las abejas ven colores. Su rechazo a admitir que los colores de las flores estaban allí sin razón alguna o simplemente para delicia de los ojos humanos, fue el estímulo para realizar estos experimentos (Dawkins 1982:31).

Un proceso deductivo similar condujo al descubrimiento de las endoríinas, sustancias similares a la morfina que producimos en nuestros propios cuerpos cuando estamos sometidos a estrés o dolor; originando, por ejemplo, ese momento de la máxima respuesta competitiva en un corredor. En el caso de las endorfinas el razonamiento fue una reversión del razonamiento de Von Frisch. Los científicos descubrieron en el cerebro unos receptores que eran extraordinariamente específicos para la morfina, sustancia que posee un potente efecto analgésico. La ingeniería revertida insiste en que donde quiera que haya una cerradura específica debe haber también una llave específica para ella. ¿Por qué existen estos receptores en el cerebro? (¡La madre naturaleza no podía haber previsto el desarrollo farmacológico de la morfina!). Bajo ciertas condiciones, deben ser algunas moléculas producidas internamente las que se comportan como llaves originales diseñadas para estas cerraduras (receptores). Busquemos una molécula que se acople a este receptor y que se produzca bajo circunstancias en las cuales una inyección de morfina pueda ser beneficiosa. ¡Eureka! ¡La morfina creada endógenamente —la endorfina— ha sido descubierta!

Incluso saltos deductivos más bien deudores del estilo de Sherlock Holmes se han ejecutado. He aquí, por ejemplo, un gran misterio: «¿Por qué

algunos genes cambian sus modelos de expresión dependiendo de si son heredados de la madre o del padre?» (Haig y Graham 1991:1.045). Este fenómeno, en el cual la maquinaria que realiza la lectura del genoma presta unas veces *más atención* al texto paterno y otras al materno, se conoce como *impresión genómica* (para una revisión general de este tema, véase Haig 1992), y su aparición ha sido confirmada en casos especiales. ¿Qué es lo que tienen en común los casos especiales? Haig y Westoby [1989] desarrollaron un modelo que se proponía resolver el misterio general mediante la *predicción* de que la impresión genómica se encontraría solamente en aquellos organismos «en los cuales las hembras tienen, a lo largo de su vida, descendencia de más de un macho y un sistema de cuidado de los hijos según el cual los hijos reciben la mayor parte de sus nutrientes de uno de los padres (generalmente la madre) y, de este modo, compite con hijos de otros machos». En estas circunstancias, razonan estos autores, existiría un conflicto entre los genes maternos y los paternos: los genes paternos tenderían a favorecer la explotación del cuerpo de la madre, tanto como fuera posible, aunque los genes maternos «verían» esta actitud casi como un suicidio, por lo que el resultado sería que los genes relevantes elegirían un bando en el tira y afloja, lo que resultaría en una impresión genómica (Haig y Graham 1991:1046).

Veamos cómo funciona el modelo de Haig. Hay una proteína, el factor de crecimiento parecido a la insulina (IGF-II) que es, como sugiere su nombre, un estimulador del crecimiento. No es sorprendente que, durante el desarrollo embrionario, las recetas genéticas de muchas especies ordenen la producción de grandes cantidades de IGF-II. Pero, como todas las máquinas en funcionamiento, el IGF-II necesita el apropiado ambiente de soporte para realizar su trabajo y, en este caso, precisa de la ayuda de moléculas conocidas como «receptores tipo 1». Hasta ahora nuestra historia es como la historia de las endoríinas: tenemos un tipo de llave (IGF-II) y un tipo de cerradura (receptores tipo 1) en la cual se acomoda aquélla y realiza, obviamente, una función importante. Pero sucede que en el ratón, por ejemplo, existe otro tipo de cerradura —los receptores tipo 2— en la cual también se acomoda la llave. ¿Para qué son estas cerraduras secundarias? Para nada, aparentemente; son descendientes de moléculas que en otras especies (los sapos, por ejemplo) desempeñan un papel en sistemas de células «eliminadoras de desechos», pero no es eso lo que hacen cuando se ligan con la IGF-II en el ratón. Entonces ¿por qué se encuentran allí? Porque, naturalmente, tal ha sido la «orden» de acuerdo con la receta genética para hacer un ratón, pero aquí está

la triquiñuela de la historia: mientras que ambas contribuciones al cromosoma, la maternal y la paternal, contienen recetas de instrucciones para hacerlo, estas instrucciones *se expresan de manera preferente* a partir del cromosoma materno. ¿Por qué? Para contrarrestar la instrucción en la receta que daría lugar a la producción en exceso del estimulador del crecimiento. Los receptores tipo 2 están precisamente allí para limpiar —para «capturar y degradar»— todos los estimuladores del crecimiento en exceso que el cromosoma paterno bombearía en el feto si no se pusiera freno. Dado que en los ratones hay especies en las que las hembras tienden a emparejarse con más de un macho, éstos compiten para explotar los recursos de cada hembra, una competición de la cual las hembras deben protegerse a sí mismas (y a sus propias contribuciones genéticas).

El modelo de Haig y Westoby predice que los genes se desenvolverían en el ratón para proteger a las hembras de su explotación y esta predicción ha sido confirmada. Además, el modelo predice que los receptores tipo 2 no deben funcionar de esta manera en las especies en las que no se plantea un conflicto genético de este tipo. No funcionan de esta manera en los pollos debido a que los hijos no pueden influir en la cuantía de la yema que reciben sus huevos, así que las guerras de prestigio nunca tendrán lugar. En cierta ocasión, Bertrand Russell describió cínicamente como argumento ilícito aquel que parece tener las ventajas de una apropiación indebida frente a un laborioso y honrado trabajo; en este sentido, uno puede simpatizar con el biólogo molecular, duro trabajador, que reacciona con cierta envidia cuando alguien como Haig se presenta bruscamente diciendo: «¡Mire debajo de esa piedra; estoy seguro de que encontrará un tesoro con la siguiente forma!».

Esto es lo que hizo Haig: predijo lo que habría sido, hace cien millones de años, el movimiento de la madre naturaleza en el juego del diseño de los mamíferos. De todos los movimientos posibles y disponibles, Haig vio que había una buena razón para este movimiento, de modo que esto fue lo que descubriría. Podemos obtener una idea acerca de la magnitud del salto que representa la inferencia de Haig, comparándolo con un salto paralelo que podemos hacer en el «Juego de la Vida». Recordemos que uno de los posibles ciudadanos del mundo de la vida es una máquina Turing Universal compuesta de trillones de *pixels*. Dado que a la máquina Turing Universal le es posible computar cualquier función computable, puede jugar al ajedrez simplemente imitando el programa que nos guste de cualquier ordenador que juegue al ajedrez. Supongamos que la máquina de Turing ocupa el «plano de la vida», jugando al ajedrez contra sí misma, como hacía el ordenador de Samuel.

Observar la configuración de los puntos que componen esta maravilla no aclararía nada a quien no tuviera la clave de la existencia de una configuración con tales poderes. Pero para alguien que *conociese la hipótesis* de que este enorme dispositivo de puntos negros era un ordenador que jugaba al ajedrez, se abriría una inmensidad de posibilidades de predecir el futuro de esa configuración.

Veamos ahora los ahorros que hemos conseguido. En un principio, estábamos enfrentados con una pantalla en la que trillones de *pixels* se encendían y se apagaban. Dado que conocemos la simple ley de la física de la vida, podríamos calcular laboriosamente, si quisiéramos, la conducta de cada punto en la pantalla, lo que tomaría su tiempo. Como un primer paso para reducir costes, dejaríamos de pensar en *pixels* individuales, para pensar en deslizadores, comilones y naturalezas muertas, y así sucesivamente. Dondequiera que viéramos un deslizador aproximándose a un comilón podríamos predecir su «desaparición en cuatro generaciones», sin necesidad de aburrirnos con cálculos a nivel de los *pixels*. En un segundo paso, podríamos pensar que los deslizadores eran símbolos en la «cinta» de una gigantesca máquina de Turing y entonces, adoptando esta posición del más elevado diseño respecto a la configuración, predecir su futuro *como* una máquina de Turing. A este nivel estaríamos «simulando manualmente» el «lenguaje de la máquina» de un programa de ordenador que juega al ajedrez, y que aun siendo una manera tediosa de hacer predicciones, sería mucho más eficiente de lo que se conseguiría mediante la física. Como tercer paso, aún más eficiente, podríamos pasar por alto los detalles del programa de ajedrez y asumir que, sean lo que sean, ¡son buenos! Es decir, podemos aceptar que el programa de ajedrez en la máquina de Turing hace que los deslizadores y los comilones no sólo jueguen ajedrez según las reglas sino buen ajedrez según las reglas; que ha sido bien diseñado (quizás ha sido diseñado por sí mismo, como el programa de damas de Samuel) para encontrar los mejores movimientos. Esto nos permite pensar en las posiciones en el tablero de ajedrez, los movimientos posibles y los fundamentos para evaluarlos, es decir, nos permite razonar acerca de las razones.

Adoptando la posición intencional acerca de la configuración podemos predecir su futuro *como* jugador de ajedrez realizando acciones intencionales: haciendo movimientos de ajedrez y tratando de conseguir un jaque mate. En primer lugar, tendríamos que deducir el esquema de interpretación que nos permita decir cuáles entre las configuraciones de *pixels* funcionan como símbolos: qué patrón de deslizadores expresa «D×A+» (Reina toma Alfil;

jaque) y los otros símbolos de los movimientos ajedrecísticos. Pero entonces podemos utilizar el esquema de interpretación para predecir, por ejemplo, que la siguiente configuración de la galaxia que aparezca sería tal y tal movimiento de un deslizador, es decir, los símbolos para «R×D» (Rey toma Reina). En todo esto hay un riesgo, ya que el programa de ajedrez ejecutado en la máquina de Turing está lejos de ser perfectamente racional y, a un nivel diferente, pueden aparecer desechos en escena y «romper» la configuración de la máquina antes de que finalice el juego. Pero si todo va bien, como normalmente sucede, si tenemos la interpretación correcta, podemos sorprender a nuestros amigos diciendo algo parecido a esto: «Mi predicción es que la próxima oleada de deslizadores que aparecerá en la localización L en esta galaxia de la vida tendrá la siguiente configuración: una configuración singular seguida de un grupo de tres, seguida de otra configuración singular». ¿Cómo puede usted predecir que aparecerá una específica configuración molecular[84]?

En otras palabras, los modelos reales pero (potencialmente) ruidosos abundan en tal configuración del mundo de la vida, para ser captados allí sólo si somos lo bastante afortunados o inteligentes para apuntar en la perspectiva correcta. No son modelos *visuales*, sino, digamos, *intelectuales*. La pantalla de un ordenador, ya sea mirándola de reojo o bien de frente, no está en condiciones de ayudarnos, mientras que la proposición de interpretaciones imaginativas (o lo que Quine denominaba «hipótesis analíticas») puede descubrir una mina de oro. La oportunidad a la que se enfrenta el observador del mundo de la vida equivale a la posibilidad que se le plantea al criptógrafo que contempla con asombro un nuevo fragmento de un texto cifrado, o la que se presenta a un marciano que esté observando a través de un telescopio la final del fútbol americano conocida como la Superbowl. Si el marciano escoge la posición intencional —también llamada psicología popular[85]— como el nivel correcto para observar el modelo, las formas emergerán fácilmente a través de los ruidosos empellones de los jugadores-partículas y de los equipos-moléculas.

Cuando uno adopta la posición intencional, la escala de comprensión con respecto a la galaxia en dos dimensiones que juega al ajedrez en el ordenador es asombrosa: ésta es la diferencia entre deducir en nuestro cerebro cuál es el movimiento más probable (mejor) de las piezas blancas en comparación con el procedimiento de calcular el estado de unos pocos trillones de *pixels* a lo largo de unos cientos de miles de generaciones. Pero la escala de los ahorros conseguidos no es realmente mayor en la escala del mundo de la vida que en

nuestro mundo. Predecir que si lanzamos un ladrillo contra alguien éste se agachará para esquivarlo es fácil desde la posición intencional de la psicología popular; por el contrario, es y será siempre espinoso tratar de seguir el rastro de los fotones desde el ladrillo al ojo, o el de los neurotransmisores desde el nervio óptico al nervio motor, y así sucesivamente.

Se puede estar preparado para pagar un exorbitante precio en errores por tan enorme poder de computación, pero, de hecho, la posición intencional, si se utiliza correctamente, proporciona un sistema de descripción que permite una predicción extraordinariamente fiable no sólo de la conducta humana inteligente, sino también de la «conducta inteligente» del proceso que ha diseñado organismos. Todo esto que estamos diciendo sería muy alentador para alguien como William Paley. Podemos trasladar el peso de la prueba a los escépticos con un simple reto argumental: si no hubiera diseño en la biosfera ¿cómo iba a *funcionar* la posición intencional? Incluso podemos obtener una medida aproximada del diseño en la biosfera si se compara el coste de hacer predicciones desde el más bajo nivel de la posición física (la cual supone que no hay diseño, o que casi no lo hay, dependiendo de cómo tratemos la evolución de los universos) con el coste de hacer predicciones desde las posiciones más elevadas: la posición del diseño y la posición intencional. La palanca añadida de la predicción, la disminución de la incertidumbre, la reducción del enorme espacio de investigación o búsqueda a unos pocos caminos óptimos o casi óptimos, es una medida del diseño observable en el mundo.

Los biólogos denominan *adaptacionismo* a este estilo de razonamiento, el cual ha sido definido por uno de sus más eminentes críticos como la «creciente tendencia en la biología evolucionista a reconstruir o predecir acontecimientos de la evolución mediante la hipótesis de que todos los caracteres se establecen en la evolución por selección natural directa del estado que ha conseguido la mejor adaptación, esto es, del estado que es una "solución" óptima para el "problema" planteado por el medio ambiente» (Lewontin 1983). Estos críticos afirman que, aunque el adaptacionismo desempeña *algún* papel importante en biología, no todo es tan central ni tan ubicuo y, naturalmente, debemos procurar un equilibrio con otros modos de pensar. He venido mostrando, sin embargo, que el adaptacionismo tiene un papel crucial en el análisis de todo suceso biológico en cualquier escala de la creación, desde la primera macromolécula autorreplicante hasta el punto más alto. Si descartamos el razonamiento adaptacionista debemos descartar, por

ejemplo, el mejor argumento de libro de texto para el propio acontecimiento de la evolución (cito la versión de Mark Ridley 1985:136): la ampliamente extendida existencia de homologías, aquellas sospechosas similitudes de diseños que *no* son funcionalmente necesarias.

El razonamiento adaptacionista no es opcional; es el corazón y el alma de la biología evolucionista. Aunque puede complementarse y sus fallos se pueden reparar, pensar en *desplazarlo* de su posición central en la biología es imaginar no ya la caída del darwinismo, sino el fracaso de la bioquímica moderna y de todas las ciencias de la vida y de la medicina. En este sentido, sorprende un poco descubrir que ésta es precisamente la interpretación que muchos lectores han dado a la más famosa e influyente crítica del adaptacionismo, el tan citado y reimpreso clásico libro de Stephen Jay Gould y Richard Lewontin, titulado *The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme* (1979) que ha sido masivamente mal interpretado.

2. *El paradigma de Leibniz*

> Si, entre todos los mundos posibles, ninguno hubiese sido mejor que el resto, entonces Dios nunca lo hubiese creado.
>
> Gottfried Wilhem Leibniz, *Teodicea*

> El estudio de la adaptación no es una preocupación opcional con fragmentos fascinantes de historia natural, sino que es el núcleo de los estudios biológicos.
>
> Colin Pittendrigh,
> *«Adaptation, Natural Selection and Behaviour»*

Es bien conocido que Leibniz dijo que éste era el mejor de los mundos posibles, una llamativa sugerencia que a distancia puede parecer presuntuosa, pero cuyo resultado ha sido, como hemos visto, arrojar una interesante luz sobre cuestiones profundas acerca de lo que significa ser un mundo posible y de lo que podemos inferir acerca del mundo real por el hecho de su realidad. Voltaire, en su obra *Candide*, creó una famosa caricatura de Leibniz, el doctor Pangloss, el erudito simplón capaz de racionalizar cualquier calamidad o deformidad —desde el terremoto de Lisboa hasta las enfermedades venéreas — y demostrar cómo, sin duda, todo esto sucedía para lo mejor. Nada, *en principio*, podía probar que este no era el mejor de rodos los mundos posibles.

Gould y Lewontin calificaron memorablemente los excesos del adaptacionismo como el «paradigma panglossiano» y lucharon para ridiculizarlo y sacarlo del escenario de la ciencia seria. No fueron los primeros en utilizar «panglossiano» como un término crítico en la teoría de la evolución. El biólogo evolucionista J. B. S. Haldane propuso una famosa lista de tres «teoremas» de deficiente argumentación científica: el teorema de Bellman («Lo que yo te diga tres veces es verdad», del libro *The Hunting of the Snark* de Lewis Carroll); el teorema de la Tía Jobisca («Es un hecho que todo el mundo conoce», del poema «The Pobble Who Had No Toes» de Edward Lear) y el teorema de Pangloss («Todo es para lo mejor en el mejor de los mundos posibles» de la novela *Candide* de Voltaire). John Maynard Smith utilizó especialmente el último de estos teoremas de Haldane para referirse a «la vieja falacia panglossiana de que la selección natural favorece las adaptaciones que son buenas para las especies en su globalidad más que actuando a nivel individual». Como más tarde comentó el propio Maynard, «es irónico que la expresión "teorema de Pangloss" fuera utilizada en primer lugar en el debate acerca de la evolución (impresa, creo, por mí mismo aunque tomada en préstamo de un comentario de Haldane), no como una crítica de las explicaciones adaptativas, sino específicamente como una crítica del "seleccionismo grupal" y de los argumentos que maximizan la capacidad media de adaptación» (Maynard Smith 1988:88). Pero Maynard Smith está en apariencia equivocado. Recientemente Gould ha llamado la atención sobre el uso aún más temprano del término por un biólogo, William Bateson (1909), uso del cual Gould no era consciente cuando decidió aplicarlo. Como Gould dice «la convergencia es difícilmente sorprendente, en cuanto que el doctor Pangloss es una sinécdoque estándar de esta forma de ridículo» (Gould 1993a:312). Como hemos visto en el capítulo 6, mientras más apta o idónea es una idea, más probable es que haya nacido (o haya sido tomada como préstamo) de modo independiente, en más de un cerebro.

Voltaire creó la figura del doctor Pangloss como una parodia de Leibniz, lo que no deja de ser exagerado e injusto con respecto a Leibniz, como sucede en toda buena parodia. Gould y Lewontin, de modo similar, hicieron la caricatura del adaptacionismo en el artículo en el que lo atacaban, de modo que la paridad del razonamiento sugiere que, si deseamos deshacer el daño de esta caricatura, y describir el adaptacionismo de un modo preciso y constructivo, tenemos a mano un título ya confeccionado: podemos denominar al adaptacionismo, considerado honestamente, como el «paradigma de Leibniz».

El artículo de Gould y Lewontin ha originado un curioso efecto en el mundo académico. Es ampliamente considerado por filósofos y otros humanistas, que han oído hablar de él o que incluso lo han leído, como una especie de *refutación del adaptacionismo*. Tuve la primera noticia de este artículo a través del filósofo-psicólogo Jerry Fodor, un crítico a lo largo de toda su vida de mi tesis acerca de la posición intencional, el cual sostenía que lo que yo estaba diciendo era puro adaptacionismo (tenía razón en este punto), y persistió hasta permitirme conocer lo que todos los conocedores saben: el artículo de Gould y Lewontin ha presentado el adaptacionismo «como si se encontrara en completa bancarrota» (véase Fodor 1990:70). Cuando me interesé por este tema, lo vi de otra manera. En 1983, publiqué un trabajo en *Behavioral and Brain Sciences* (BBS) titulado «Intentional Systems in Cognitive Ethology», y dado que era desenfadadamente adaptacionista en su razonamiento, incluí una coda titulada «La defensa del paradigma panglossiano» que criticaba tanto a Gould como a Lewontin y, especialmente, el extraño mito que había surgido alrededor de este tema.

Los resultados fueron fascinantes. Todo artículo que apareciera en BBS venía acompañado de varias docenas de comentarios de expertos en campos relevantes y mi pieza atrajo el fuego de biólogos evolucionistas, psicólogos, etólogos y filósofos, la mayoría de ellos en tono amistoso aunque algunos eran extraordinariamente hostiles. Una cosa era clara: no eran precisamente algunos filósofos y psicólogos quienes se encontraban incómodos con el razonamiento adaptacionista. Además de los evolucionistas teóricos que apostaban con entusiasmo a mi favor (Dawkins 1983b, Maynard Smith 1983), y aquellos que estaban en contra (Lewontin 1983), había otros que, aunque estaban de acuerdo conmigo en que Gould y Lewontin no habían refutado el adaptacionismo, eran partidarios de minimizar el uso estándar de las suposiciones de «optimalidad» que yo defendía como un ingrediente esencial de todo el pensamiento evolucionista.

Niles Eldredge discutió la ingeniería revertida de los morfólogos funcionales: «Lograréis soberbios análisis de puntos de apoyo, vectores de fuerzas y así en adelante, hasta llegar a entender la anatomía como una máquina viviente. Alguno de estos elementos es muy bueno. Alguno es absolutamente espantoso» (1983:361). Eldredge citaba, como ejemplo de buena ingeniería revertida, el trabajo de Dan Fisher [1975] en el que comparaba los modernos cangrejos en herradura con sus ancestros jurásicos:

Asumiendo solamente que los cangrejos jurásicos en herradura también nadaban sobre sus dorsos, Fisher demostró que debían nadar en un ángulo de o-10 grados (a plano sobre sus dorsos) a una velocidad algo mayor de 15-20 cm/sec. Así la «significación adaptan va» de las leves diferencias en anatomía entre los modernos cangrejos en herradura y sus parientes de hace 150 millones de años se traduce en una comprensión de sus capacidades natatorias ligeramente diferentes. (Con toda honestidad, debo también recordar que Fisher utiliza el término «optimalidad» en sus argumentos. Fisher entiende las diferencias entre las dos especies como una especie de compensación beneficiosa, donde los nadadores jurásicos ligeramente más eficientes parecen haber utilizado las mismas piezas anatómicas para excavar con menor eficiencia que sus parientes modernos). En cualquier caso, el trabajo de Fisher se presenta como un ejemplo realmente bueno de análisis morfológico funcional. La noción de adaptación no es más que una filigrana conceptual que pudo haber desempeñado un papel en la motivación de la investigación, pero que no era vital para la propia investigación (Eldredge 1983:362).

Pero de hecho el papel de las suposiciones de optimalidad en el trabajo de Fisher —más allá del papel explícito que le concede Eldredge— es tan «vital» y naturalmente tan omnipresente que Eldredge lo pasó por alto completamente. Por ejemplo, la inferencia de Fisher de que los cangrejos jurásicos nadaban 15-20 cm/sec tiene como premisa tácita que aquellos cangrejos *nadaban a la velocidad óptima para su diseño*. (¿Cómo sabe Fisher que nadaban? Quizá los cangrejos yacían allí, ajenos a la excesiva funcionalidad de sus formas corporales). Sin esta tácita (y completamente obvia) premisa, no puede extraerse conclusión alguna acerca de lo que sería la velocidad de natación real de la variedad jurásica de los cangrejos en herradura.

Michel Ghiselin fue más directo en la negación de esta obvia dependencia, nada obvia:

El panglossianismo es malo debido a que plantea la pregunta errónea, es decir, «¿Qué es bueno?». La alternativa es rechazar totalmente esta teleología. En lugar de preguntar «¿Qué es bueno?», nosotros preguntamos «¿Qué ha sucedido?». La nueva pregunta abarca lo mismo que la antigua y mucho más (1983:363).

Ghiselin se confundía. Difícilmente hay una simple respuesta a la pregunta «¿Qué ha sucedido (en la biosfera)?» que no dependa crucialmente de las suposiciones acerca de lo que es bueno[86]. Como hemos señalado, ni siquiera podemos avalar el concepto de homología sin apoyarnos en el adaptacionismo y sin hacer lo mismo con la posición intencional.

¿Cuál es ahora el problema? El problema es cómo distinguir el buen —e irreemplazable— adaptacionismo del malo, cómo distinguir a Leibniz de Pangloss[87]. Seguramente una razón para la extraordinaria influencia entre los no evolucionistas del artículo de Gould y Lewontin es que expresa, con muchas bellas florituras retóricas, lo que Eldredge llamó «la resaca» contra el concepto de adaptacionismo entre los biólogos. ¿Contra quiénes reaccionaban? Fundamentalmente contra cierto tipo de pereza: contra los adaptacionistas que dieron con una explicación excelente y verdadera mediante la cual una determinada circunstancia debía prevalecer, y a partir de ese momento nunca se molestaron en someterla a prueba, dado que presumían que era una historia demasiado buena para no ser verdadera. Adoptando otra etiqueta literaria, en esta ocasión de Rudyard Kipling [1912], Gould y Lewontin calificaron esta explicación de ‘Al fin y al cabo, historias’. Es una curiosidad histórica que Kipling escribiera su obra *Just So Stories* en un tiempo en el que esta objeción a las explicaciones darwinianas venía dando vueltas desde hacía varias décadas[88]; algunas de éstas fueron traídas a colación por algunos de los primeros críticos de Darwin (Kitcher 1985a:156). ¿Se inspiró Kipling en esta controversia? En cualquier caso, considerar las elucubraciones de los adaptacionistas como «historias, al fin y al cabo» difícilmente los acredita; aunque he encontrado siempre deliciosas las fantasías de Kipling acerca de cómo el elefante consiguió su trompa, y el leopardo sus manchas, se trata en realidad de cuentos muy simples y nada sorprendentes, cuando se los compara con las atractivas hipótesis que han elaborado los adaptacionistas.

Consideremos el caso del *Prodotiscus insignis*, pájaro de los bosques de África, conocido vulgarmente como «honey guide» (literalmente ‘guía para la miel’), en su variedad más grande (*Indicator indicator*), el cual debe el calificativo a su talento para guiar a los seres humanos hasta las colmenas salvajes ocultas en el bosque. Cuando los boran, en Kenia, desean encontrar miel, llaman al pájaro soplando silbatos especialmente fabricados con conchas de caracoles. Cuando llega uno de estos pájaros, vuela alrededor cantando una especial canción denominada «sígueme». Los boran siguen al

pájaro que vuela rápidamente delante y los espera hasta ser alcanzado, asegurándose continuamente de que lo pueden ver cuando vuela. Una vez que el pájaro ha llegado a la colmena, sustituye la melodía de su canto inicial por otra que viene a decir «aquí hemos llegado». Cuando los boran localizan la colmena en el árbol y la abren, extraen la miel, dejando la cera y las larvas para el pájaro «guía de la miel». Ante esta historia, ¿no estamos dispuestos a creer que existe realmente esta admirable cooperación y que posee las inteligentes propiedades funcionales descritas? ¿No estamos dispuestos a aceptar que tal maravilla puede haber evolucionado bajo alguna imaginada serie de presiones de selección y de oportunidades? Yo ciertamente sí estoy dispuesto. Y en este caso, felizmente, la investigación que ha llevado a cabo el seguimiento está confirmando la historia, e incluso ha añadido, mientras progresaba, excelentes detalles. Recientes pruebas controladas, por ejemplo, demuestran que los buscadores de miel de la tribu boran tardan mucho más tiempo en encontrar las colmenas sin la ayuda de los pájaros y que el 96 por ciento de las 180 colmenas halladas durante el estudio estaban en árboles situados en lugares que hubieran sido inaccesibles para los pájaros sin la ayuda de los seres humanos (Isack y Reyer 1989).

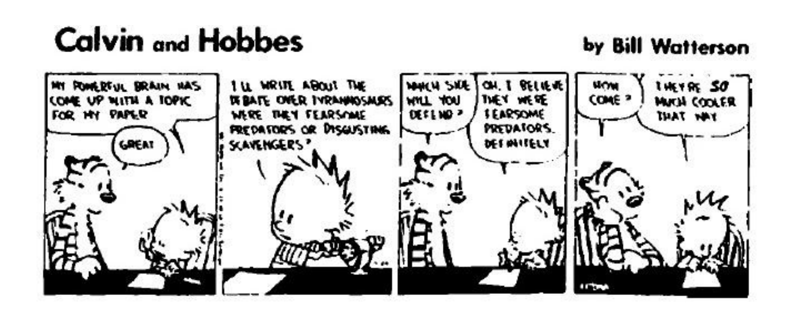


Figura 9.3

Otra fascinante historia, más cercana a nosotros, es la hipótesis de que nuestra especie, el *Homo sapiens*, ¡desciende de los primates más precoces, a través de una vía de especies intermedias que fueron acuáticas! (Hardy 1960, Morgan 1982, 1990). Estos simios acuáticos habrían vivido presuntamente en las playas de una isla formada por la inundación del área que ahora es Etiopía, durante el último período Mioceno, hace alrededor de siete millones de años. Separados por las aguas de sus primos del continente africano y bajo el reto de un cambio relativamente súbito del clima y de las fuentes de alimentos, desarrollaron el gusto por los crustáceos, y durante un período de más o menos un millón de años iniciaron el proceso evolutivo de retornar al mar,

proceso al que fueron sometidos previamente, por ejemplo, ballenas, delfines, focas, nutrias y otros animales. Estaba en marcha este proceso, que condujo a la fijación de muchas características curiosas que, por otro lado, se encuentran *solamente* en los mamíferos acuáticos —no en otros primates— cuando las circunstancias cambiaron de nuevo y estos monos semiacuáticos retornaron a la vida terrestre (aunque al borde del mar, de los lagos o de los ríos). Allí, estos simios acuáticos descubrieron que muchas de las adaptaciones que, por buenas razones, habían desarrollado en los días en los que buceaban para coger crustáceos, ahora no sólo no eran valiosas sino evidentes estorbos. No obstante, pronto consiguieron reconvertir esas dificultades en utilidades o, al menos, las compensaron: la posición erecta, la postura bipedal, la capa subcutánea de grasa, la lampiñez, la transpiración, las lágrimas, la incapacidad para responder a la privación de sal según el modelo estándar de los mamíferos y, naturalmente, el reflejo de inmersión (que permite incluso a los recién nacidos humanos sobrevivir a la inmersión súbita en el agua durante largos períodos, sin efectos nocivos). Los detalles —y hay muchos, muchos más— son tan ingeniosos, y la globalidad de la teoría del simio-acuático choca de tal manera con lo establecido, que personalmente me encantaría verla justificada. Esto no la convierte en verdadera, naturalmente.

El hecho de que en estos días la principal exponente de esta tesis de los monos acuáticos sea no sólo una mujer, Elaine Morgan, sino una aficionada, una escritora científica sin credenciales propiamente oficiales, a pesar de sus sustanciales investigaciones, hace que la perspectiva de su justificación sea extraordinariamente atrayente[89]. El *establishment* ha respondido con gran ferocidad a sus retos, tratándolos en su mayoría como una noticia que no merecía la menor consideración, aunque ocasionalmente sometiéndola a un rechazo displicente[90]. Esta no es, necesariamente, una reacción patológica. La mayoría de los que proponen «revoluciones» científicas son excéntricos sin credenciales que no merecen que se les preste atención. Existe realmente un grupo de estos excéntricos que nos asedia continuamente, y la vida es demasiado corta para tener que dedicar tiempo a reflexionar sobre cada hipótesis que no ha sido solicitada. Pero, en este caso, me sorprendo ya que muchas de las refutaciones me parecen terriblemente débiles y *ad hoc*. En los últimos años, cada vez que me encontraba en compañía de distinguidos biólogos, teóricos de la evolución, paleoantropólogos y otros expertos, a menudo les rogaba que me dijeran exactamente por qué Elaine Morgan estaba equivocada con su teoría del mono acuático. Todavía no he conseguido una respuesta digna de mención, sin contar a quienes admiten, con un guiño, que a

menudo se han formulado la misma pregunta. Para éstos no parece existir en la idea nada inherentemente imposible; después de todo, otros mamíferos se han dado la zambullida. ¿Por qué no podían nuestros antepasados haber comenzado volviendo al océano para después retornar, trayendo consigo algunas cicatrices, como indicios de esta historia?

A Morgan se le puede «acusar» de contar una buena historia —ciertamente lo es— pero no de haberse negado a someterla a prueba. Por el contrario, Elaine Morgan ha utilizado la historia como una palanca para conseguir innumerables y sorprendentes predicciones en diversos campos, y ha estado dispuesta a ajustar su teoría cuando los resultados lo exigían. Por otro lado, Elaine ha disparado sus armas y, de hecho, ha invitado al ataque de sus puntos de vista, debido a la vehemencia de su partidismo. Como sucede a menudo en estos enfrentamientos, la intransigencia y la actitud defensiva, por ambos lados, ha comenzado a pasar factura, dando lugar a uno de esos espectáculos que descorazonan a cualquiera que desea conocer la verdad acerca de una cuestión. No obstante, el último libro de Morgan [1990] sobre este tema responde con admirable claridad a las objeciones que se habían acumulado hasta la fecha, y útilmente contrasta las fuerzas y las debilidades de la teoría del mono acuático con las pertenecientes a la historia del *establishment*. Y más recientemente aún, ha aparecido un libro que recoge ensayos de diversos expertos, a favor y en contra de la teoría del mono acuático. El veredicto provisional de los organizadores de la conferencia de 1987 de la cual nace este libro fue que «mientras que hay un cierto número de argumentos que favorecen la AAT, éstos no son lo suficientemente convincentes para contrarrestar los argumentos contra la teoría» (Roede y otros 1991:324). Esta juiciosa nota, que muestra un moderado desacuerdo con Morgan, asegura que la discusión continuará, quizás incluso con menos rencor; será interesante ver dónde terminará todo esto.

Mi objetivo al plantear la teoría del mono acuático no es defenderla del punto de vista del *establishment* biológico, sino utilizarla como ilustración de una preocupación más profunda. A muchos biólogos les gustaría decir: «¡Qué pesadas son ambas partes!». Morgan [1990] expone con habilidad el pensamiento voluntarista y apasionado que ha calado en el relato del *establishment* acerca de cómo —y *por qué*— el *Homo sapiens* desarrolló la bipedestación, la sudoración y la piel lampiña sobre la sabana y no a orillas del mar. En comparación con la historia de Elaine Morgan, las historias del *establishment* pueden no ser literalmente tan inverosímiles, aunque algunas sí que son muy improbables, incluso igual de especulativas y (me aventuro a

decirlo) no mejor confirmadas. Lo que los miembros del *establishment* han procurado, hasta donde yo sé, es ocupar el lugar más preeminente en los libros de texto, antes de que Hardy y Morgan trataran de desalojarlos. Ambas partes han sido indulgentes con el adaptacionismo, considerado como «historias, al fin y al cabo» y, dado que *una u otra historia* puede ser verdadera, no debemos concluir que hemos encontrado *la* historia justa porque nos hayamos encontrado con *una* en la que parecen encajar los hechos. Lo que sí es un exceso que merece ser criticado en los adaptacionistas es que hayan sido menos activos en la búsqueda de confirmaciones (o definitivos rechazos) de sus historias[91].

Pero antes de abandonar esta cuestión, deseo subrayar que hay muchas historias adaptacionistas que todo el mundo acepta con gusto, aun cuando no hayan sido nunca «sometidas a prueba de manera apropiada», precisamente porque son tan obviamente verdaderas como para merecer posteriores comprobaciones. ¿Duda alguien que los párpados han evolucionado para proteger los ojos? Pero esta clarísima obviedad puede ocultar buenas cuestiones para investigar. Georges Williams señala que, ocultos tras hechos obvios, pueden encontrarse otros que merecen ser investigados:

> El parpadeo de un ojo humano dura 50 milisegundos. Esto significa que estamos ciegos alrededor del 5 por ciento del tiempo cuando utilizamos los ojos normalmente. Muchas cosas importantes pueden suceder en 50 milisegundos, de modo que podemos perderlas de vista totalmente. Una piedra o una lanza arrojada por un poderoso adversario pueden desplazarse más de un metro en 50 milisegundos, por lo que sería importante percibir este movimiento con tanta precisión como fuera posible. ¿Porqué parpadeamos con ambos ojos simultáneamente? ¿Porqué no lo hacemos alternativamente y sustituimos nuestra atención visual del 95 por ciento por el cien por cien? Puedo imaginar una respuesta basada en algún tipo de análisis «coste-beneficio», Un mecanismo de parpadeo simultáneo de ambos ojos es mucho más simple y más barato que un parpadeo alternativo (G. Williams 1992:152-153).

Williams no ha intentado hasta ahora confirmar o rechazar cualquier hipótesis a partir de esta pieza ejemplar de planteamiento adaptacionista de un problema concreto, aunque ha estimulado la investigación, planteando

preguntas. Sería un ejercicio tan puro de ingeniería revertida como pudiera imaginarse.

Serias consideraciones acerca de por qué la selección natural permite el parpadeo simultáneo pueden generar, por otra parte, cuestiones cuya explicación se nos escapa. ¿Qué cambios serían necesarios en la maquinaria para producir el primer paso hacia una imaginada alternancia adaptativa o simplemente un parpadeo no simultáneo? ¿Cómo puede conseguirse el cambio desde el punto de vista del desarrollo? ¿Qué otro tipo de cambios pueden esperarse de una mutación que produzca un ligero intervalo en el parpadeo de un ojo respecto al otro? ¿Cómo actuaría la selección sobre tal mutación? (G. Williams 1992:153).

El propio Gould ha apoyado algunas de las más provocativas y sugerentes historias sobre adaptaciones («Historias, al fin y al cabo»), tales como la argumentación de Lloyd y Dybass [1966] para explicar por qué las cigarras (las denominadas «cigarras periódicas» o «cigarras con ciclos vitales de diecisiete años») tienen ciclos reproductivos que duran un número primo de años; trece o diecisiete, pero nunca quince ni dieciséis, por ejemplo. «Como evolucionistas», dice Gould, «buscamos respuestas a esta cuestión. Concretamente, ¿por qué se ha evolucionado hasta llegar a esta sorprendente sincronía, y por qué debe ser tan largo el período que transcurra entre los episodios de reproducción sexual?» (Gould 1977a:99)**[92]**. La respuesta —que posee, retrospectivamente, un bello sentido— es que, al existir un elevado número primo de años entre sus apariciones, las cigarras disminuyen la posibilidad de que sus predadores las descubran y las persigan, predadores que aparecen cada dos, tres o cinco años. Si las cigarras tuvieran una periodicidad, por ejemplo, de dieciséis años, entonces serían un raro bocado para los predadores que apareciesen cada año, pero una más fiable fuente de alimento para los predadores que apareciesen cada dos o cuatro años, y una apuesta segura para los predadores que entrasen en fase con las cigarras en períodos de ocho años. Sin embargo, si su período no es múltiplo de ningún número más bajo, la posibilidad de servir de alimento es tan rara que no merece la pena de ser seguida por ninguna especie que no fuera lo bastante afortunada como para tener exactamente su periodicidad o algún múltiplo de ella. No sé si las historias de Lloyd y Dybass («historias, al fin y al cabo») han sido apropiadamente confirmadas, pero no creo que Gould sea culpable de panglossianismo al considerarlas establecidas hasta que no se demuestre lo contrario. Y si realmente lo que desea Gould es preguntar y contestar preguntas, no tiene otra elección que comportarse como un adaptacionista.

Para Gould y Lewontin el problema es que no hay parámetros para medir cuándo un fragmento concreto de un razonamiento adaptacionista es algo más

que una buena idea. ¿Hasta qué punto, realmente, este problema es serio, aunque no tenga una «solución» asentada en principios? Darwin nos ha enseñado a no buscar las esencias, trazando líneas divisorias entre la función *genuina* y la intencionalidad *genuina*, y simplemente *en-su-camino-para-ser* función o intencionalidad. Cometemos un error fundamental si creemos que para aceptar el pensamiento adaptacionista necesitamos una licencia y que sólo las personas con licencia están en posesión de una estricta definición o de los criterios para una adaptación genuina. Disponemos de buenas reglas nacidas del sentido común que la ingeniería revertida prospectiva puede aplicar, las cuales hizo explícitas Georges Williams [1966]. 1) No invoquemos la adaptación mientras existan otras explicaciones, de un nivel inferior, a nuestra disposición (como las explicaciones físicas). No debemos preguntar qué ventajas se acumulan en el arce para explicar la tendencia de sus hojas a *caer*, del mismo modo que los ingenieros que se dedican a la ingeniería revertida en la Raytheon no necesitan buscar una razón que explique por qué la General Electric fabrica sus aparatos de tal modo que se funden rápidamente en los hornos, 2) No invoquemos la adaptación cuando una característica sea el resultado de alguna necesidad general del desarrollo. No necesitamos una razón especial de incrementada adaptación al medio para explicar el hecho de que las cabezas estén unidas a los cuerpos, o las piernas estén dispuestas en pares, del mismo modo que los ingenieros de la Raytheon no necesitan explicar por qué los componentes de los artefactos de la General Electric tienen tantos bordes y esquinas con ángulos rectos. 3) No invoquemos la adaptación cuando una característica es un producto colateral de otra adaptación. No necesitamos dar una explicación adaptacionista de la capacidad de los picos de los pájaros para asear su plumaje (dado que las características de los picos de los pájaros responden a razones más importantes), del mismo modo que no necesitamos explicar de manera especial la capacidad de las envolturas de los artefactos de la GE para proteger su contenido de los rayos ultravioletas.

Pero ya hemos aprendido que en cada caso estas reglas basadas en el sentido común pueden ser sobrepasadas por una investigación más ambiciosa. Supongamos que alguien maravillado por la brillantez de las hojas de los árboles en el otoño de Nueva Inglaterra se pregunta por qué las hojas de los arces brillan con tanto colorido en octubre. ¿No está este adaptacionismo desbocándose? ¿Sombras del doctor Pangloss? Las hojas brillan con esos colores sencillamente porque una vez agotada la energía acumulada durante la estación del verano, la clorofila desaparece de las hojas y las moléculas

residuales tienen propiedades reflectivas que son las que determinan los colores brillantes; una explicación a nivel de la química o de la física, no un propósito biológico. Pero ¡seamos pacientes! Aunque puede que ésta haya sido la única explicación correcta hasta ahora, actualmente es verdad que los seres humanos aprecian tanto estas hojas otoñales (el norte de Nueva Inglaterra recauda anualmente millones de dólares procedentes de los turistas) que son los turistas quienes protegen los brillantes árboles del otoño. Podemos estar seguros de que si fuéramos un árbol que estuviera luchando por sobrevivir en Nueva Inglaterra tendríamos una ventaja selectiva en caso de que nuestro ramaje fuera brillante en otoño. Aunque esto pueda ser poca cosa, y a largo plazo no represente demasiado (a largo plazo puede que no haya árboles en Nueva Inglaterra, por una u otra razón), así es, después de todo, como comienza la adaptación, como efectos fortuitos aceptados en el medio ambiente, de modo oportunista, por las fuerzas selectivas. Y, naturalmente, también hay una explicación adaptacionista para el hecho de que los ángulos rectos predominen en los productos manufacturados, y para la simetría que predomina en las extremidades orgánicas. Estas pueden llegar a ser tradiciones absolutamente establecidas, casi imposibles de suprimir por innovación, pero las razones por las cuales esas son las tradiciones no son difíciles de hallar o resultan controvertidas.

La investigación adaptacionista siempre deja cuestiones sin respuestas, abiertas para la próxima sesión. Veamos el caso de la tortuga gigante de mar y sus huevos:

Cerca del final de la puesta de huevos, es depositado un número variable de huevos pequeños, a veces deformes, que no contienen ni embrión ni yema (sólo albúmina). El propósito de estos huevos deformes no se comprende bien, aunque llegan a desecarse en el curso de la incubación y pueden moderar la humedad o el volumen de aire en la cámara incubadora. (Es posible que no cumplan función alguna o que sean el vestigio de algún mecanismo que no es evidente para nosotros en la actualidad.) (Eckert 199:30).

Pero ¿adónde conduce todo esto? Esta curiosidad sin fin del adaptacionista irrita a mucho teóricos, quienes desearían que hubiese códigos de conducta más estrictos en este campo de la ciencia. Muchos de los que han tenido la esperanza de contribuir a clarificar la controversia entre el adaptacionismo y su reacción contraria han desesperado de encontrar tales códigos, después de gastar bastante energía en diseñar y criticar varias propuestas «legislativas». No han sido suficientemente darvinianos en su modo de pensar. Mientras mejor elimina a sus rivales el pensamiento adaptacionista, por los caminos

normales, justamente una ingeniería revertida de segundo orden se traiciona a sí misma, más pronto o más tarde.

La cara del esquimal, considerada a veces como «modelada por el frío» (Coon y otros 1950), llegaría a ser una adaptación para generar y mantener potentes fuerzas masticatorias (Shea 1977). No combatimos estas nuevas interpretaciones; pueden que sean correctas. No obstante, nos preguntamos si el fallo de una explicación adaptativa debe siempre incitar a la búsqueda de otra de la misma forma considerada en su globalidad, más que proceder a una consideración de alternativas a la proposición de que cada parte es para algún propósito específico (Gould y Lewontin 1979:152).

¿Es el ascenso y la caída de sucesivas explicaciones adaptativas de diversas cosas un signo de una ciencia sana constantemente mejorando su visión o es como la patológica tergiversación de sus historias que padece el mentiroso compulsivo? Si Gould y Lewontin ofrecen una alternativa seria al adaptacionismo, la defensa de su caso antes del veredicto final será más persuasiva, pero aunque ellos y otros han luchado con energía y promovido sus alternativas con audacia, ninguna de estas alternativas ha arraigado todavía.

El adaptacionismo, el paradigma que considera los organismos como complejas máquinas adaptativas, cuyas partes poseen funciones adaptativas subsidiarias de la aptitud del conjunto, es hoy tan básico para la biología como la teoría atómica lo es para la química. Y también casi tan controvertido. Los explícitos planteamientos adaptacionistas van ascendiendo en las ciencias de la ecología, la etología y la evolución, debido a que han probado que son esenciales para los descubrimientos; si alguien duda de esta afirmación, que lea las revistas. La proclama de Gould y Lewontin en favor de un paradigma alternativo ha fracasado en su intento de impresionar a los biólogos prácticos, debido a que el adaptacionismo tiene éxito y se encuentra bien fundamentado, y a causa también de que sus críticos no disponen de un programa alternativo de investigación. Todos los años asistimos a la aparición de nuevas revistas como *Functional Biology* y *Behavioral Ecology*. Una investigación capaz de llenar el primer número de *Dialectical Biology* está todavía por materializarse (Daly 1991:219).

Lo que enfurece especialmente a Gould y a Levvontin, como sugiere el pasaje antes citado sobre la cara del esquimal, es la despreocupada confianza con que los adaptacionistas se entregan a su ingeniería revertida, siempre seguros de que más pronto o más tarde encontrarán *la razón* que explique por qué las cosas son como son, aunque se les escape durante mucho tiempo. He aquí un ejemplo, sacado de la discusión de Richard Dawkins sobre el curioso caso de los peces planos (rodaballos y lenguados, por ejemplo) que cuando nacen son peces verticales como el arenque o el pez luna, pero cuyos cráneos sufren una extraña torsión, haciendo que un ojo se acerque al otro, convirtiendo ese lado en la parte superior de ese pez que mora en las profundidades. ¿Por qué no han evolucionado igual que otros habitantes de las profundidades, que no se apoyan en uno de sus lados sino en su abdomen «como tiburones que

hubiesen pasado bajo un rodillo de vapor»? (Dawkins 1986a:91). Dawkins *imagina* una escenificación:

Aunque la vía seguida por la raya para llegar a ser un pez plano podría, en último término, haber sido también el mejor diseño, los posibles estados intermedios que se encuentran a lo largo de este camino evolutivo los recorrieron aparentemente menos bien, a corto plazo, que sus rivales que yacen en el fondo sobre uno de sus lados. Los rivales que yacen sobre un lado estaban en mejores condiciones, a corto plazo, de esconderse en el fondo del mar. En el hiperespacio genético existe una suave trayectoria que conecta los peces que nadan libremente con los peces planos que yacen sobre uno de sus lados con los cráneos retorcidos. No existe una suave trayectoria que conecte estos peces ancestrales con los peces planos que yacen en el fondo sobre su vientre. En teoría existiría trayectoria evolutiva, pero pasaría a través de estados intermedios que, de haber existido, no habrían tenido éxito, a corto plazo, que es lo único que cuenta (Dawkins 1986a:92-93).

¿Realmente lo sabía Dawkins? ¿Sabía que los estados intermedios postulados eran los menos adaptados? No lo sabía, ya que había utilizado algunos datos extraídos de fósiles. Se trata simplemente de una explicación teórica, argumentada *a priori* bajo la presunción de que la selección natural nos cuenta la verdadera historia —alguna historia verdadera u otra— acerca de cada característica curiosa de la biosfera. ¿Es esto objetable? Se da por probada la cuestión; pero ¡qué cuestión se da por probada! Se asume que el darwinismo está básicamente en el camino correcto. (Los meteorólogos dicen —dando por probada y resuelta la cuestión respecto a las fuerzas sobrenaturales— que debe haber una explicación puramente física para el nacimiento de los huracanes, aunque no se conozcan aún muchos detalles. ¿Y se puede oponer reparo a esa opinión?). Nótese que en este ejemplo la explicación de Dawkins es casi correcta, ya que no hay nada especialmente audaz en esta especulación concreta. Por otra parte, este es exactamente el tipo de reflexión que utiliza un buen técnico en la ingeniería revertida. «Parece obvio que la cubierta de este artefacto de la General Electric debe fabricarse con dos piezas en vez de tres, pero está hecha con tres piezas, lo que no deja de ser un despilfarro y, al mismo tiempo, un sistema propicio a que se produzcan filtraciones, así que podemos estar seguros de que para algunos —por cortos de vista que fueran— tres piezas era una solución mejor que dos. ¡Seguir observando!». El filósofo de la biología Kim Sterelny, en una revisión del libro *El relojero ciego* (1986), enfocó la cuestión de este modo:

Dawkins está presentando únicamente escenificaciones: mostrando que es *concebible* que (por ejemplo) las alas evolucionen gradualmente bajo la selección natural. Aun así, se puede objetar. ¿Es realmente cierto que la selección natural es de tal sutileza que el protoaguijón de un insecto que tenga un 5 por ciento de un aguijón completo es mejor que otro que tenga sólo un 4 por ciento? (1986a:82-83). Una dificultad como ésta es especialmente importante ya que los escenarios adaptativos de Dawkins no

hacen mención de los costes de estos cambios que se alegan como adaptativos. El mimetismo puede desilusionar a las parejas potenciales así como a los potenciales predadores… Creo que esta objeción es un subterfugio debido a que esencialmente estoy de acuerdo con que la selección natural es la única explicación posible para la complejidad de la adaptación. De modo que algo parecido a las historias de Dawkins ha de ser la explicación correcta (Sterelny 1988:424)[93].

3. *Jugango con limitaciones*

> Es tan absurdo quejarse de que la gente es egoísta y traidora como quejarse de que el campo magnético no aumenta a menos que el campo eléctrico tenga una disposición en espiral.
>
> John von Neumann
> (Citado en William Poundstone, *Prisoner's Dilemma*)

> Por regla general, un biólogo que observa hoy a un animal haciendo algo para beneficiar a otro, asume que éste ha sido manipulado por el otro individuo o que se comporta sutilmente como un egoísta.
>
> George Williams,
> *«Huxley's Evolution and Ethics in Sociobiological Perspective»*

Sin embargo, uno puede mostrarse razonablemente inquieto ante la amplitud del papel que alcanza en el pensamiento adaptacionista la imaginación absoluta y sin trabas. ¿Qué cabe decir con respecto a mariposas dotadas de pequeñas armas automáticas para su autoprotección? A menudo se cita este fantástico ejemplo como el tipo de opción que puede ser descartado, sin un análisis detallado, por los adaptacionistas que tratan de describir el conjunto de las posibles adaptaciones de la mariposa para las cuales la madre naturaleza ha escogido la mejor, tras haber considerado todas las opciones. Es evidente que se trata de una posibilidad demasiado distante en el «espacio de diseño» como para ser tomada en serio. Pero, como Richard Lewontin acertadamente señala, «mi sospecha es que si nunca se han hallado hormigas que cultiven hongos, la sugerencia de que ésta era una posibilidad razonable dentro de la evolución de la hormiga habría sido considerada como una estupidez» (1987:156). Los adaptacionistas son maestros en el razonamiento retrospectivo, como el jugador de ajedrez que sólo cae en la cuenta de que acaba de forzar el jaque mate en dos jugadas *después* de haber realizado el movimiento. «¡Qué jugada tan brillante! ¡Estuve muy cerca de esta idea!». Pero antes de que nuestra conclusión sea que éste es un *defecto* del método adaptacionista, debemos recordar que este reconocimiento retrospectivo de brillantez —en la jugada— es el modo según el cual siempre actúa la madre

naturaleza. Difícilmente puede culparse a los adaptacionistas de ser incapaces de predecir los brillantes movimientos que ni siquiera la madre naturaleza había tenido en cuenta hasta tropezar con ellos.

En el adaptacionismo, la perspectiva del *juego* es ubicua y la *teoría matemática de los juegos* ha desempeñado un creciente papel desde su introducción por John Maynard Smith [1972, 1974] en la teoría de la evolución[94] La teoría de los juegos de John von Neumann es una contribución fundamental al pensamiento del siglo.[95]. Von Neumann creó la teoría de los juegos en colaboración con el economista Oskar Morgenstern, y su desarrollo ha derivado de su conocimiento de que los *agentes* marcan una diferencia fundamental para la complejidad del mundo[96]. Considerando que un *agente* solitario a la manera de un Robinson Crusoe, puede ver todos los problemas como buscando los máximos estables —ascendiendo a la colina o al Monte Fuji, si se prefiere—, en cuanto otros *agentes* (que buscan máximos) son incluidos en el medio ambiente, se necesitan métodos de análisis completamente diferentes:

Un principio guía no puede ser formulado debido a la necesidad de maximizar, al mismo tiempo, dos (o más) funciones… Nos podría inducir a creer, erróneamente, que esto podría obviarse… simplemente recurriendo a la teoría de la probabilidad. Todo participante puede determinar las variables con las cuales describe sus propias acciones pero no las acciones de los otros. No obstante, desde este punto de vista aquellas variables «extrañas» no pueden describirse por medio de conjeturas estadísticas. Esto es así porque los otros (agentes) se orientan, como él mismo, por principios racionales —sea lo que sea lo que estos principios signifiquen— y ningún *modus procedendi* puede ser correcto si no intenta comprender aquellos principios y las interacciones, entre todos los participantes, de los conflictos de interés (Von Neumann y Morgenstern 1944:11).

El punto de vista fundamental que une la teoría de los juegos con la teoría de la evolución es que los «principios racionales y sea lo que sea lo que éstos signifiquen», que «guían» a los *agentes* en competición pueden ejercer su influencia incluso sobre *semi-agentes* inconscientes o irreflexivos como virus, árboles e insectos, debido a que las posibilidades de riesgo y de resultados negativos o positivos de la competición determinan qué líneas de juego, en caso de ser adoptadas, no pueden ayudar a ganar o perder, aunque se adopten de manera insensata.

El ejemplo mejor conocido en la teoría de los juegos es el *Dilema del Prisionero*, un sencillo «juego» entre dos personas, el cual arroja sombras,

obvias y sorprendentes, sobre muchas y diferentes circunstancias de nuestro mundo. Este es, básicamente, el dilema descrito (una excelente y detallada discusión se encuentra en Poundstone 1992 y en Dawkins 1989a): una persona y yo hemos sido encarcelados pendientes de juicio (con una acusación de fraude, según se dice) y el fiscal ofrece a cada uno de los prisioneros, por separado, el mismo trato: si ambos se resisten, y no confiesan ni implican al otro, cada uno recibirá una sentencia de corta duración (la evidencia del fiscal del Estado no es fuerte); si yo confieso e implico al otro y éste se resiste a hablar, saldré libre y él permanecerá de por vida en prisión; si ambos confesamos e implicamos al otro, ambos conseguiremos sentencias de prisión de duración media. Naturalmente, si permanezco callado y el otro confiesa, él saldrá libre y yo seré el condenado a cadena perpetua. ¿Qué debo hacer?

Si ambos nos resistimos a confesar, desafiando al fiscal, será mucho mejor para los dos que si ambos confesamos, así que ¿por qué no nos comprometemos los dos a permanecer callados? (En la jerga estándar del *Dilema del Prisionero* la opción de permanecer sin hablar se denomina *cooperación*). Ambos podemos llegar a este compromiso, pero cada cual puede sentir la tentación —sucumba o no a ella— de *traicionar* el compromiso, dado que, de este modo, conseguiría una libertad bajo fianza, dejando al *incauto*, triste es decirlo, en una difícil situación. Dado que el juego es simétrico, la otra persona también estará tentada, naturalmente, de hacer que yo sea el que quede como un incauto, al ser traicionado. ¿Puedo arriesgarme, manteniendo la promesa, a una sentencia de prisión de por vida? Probablemente sería más seguro *traicionar*, ¿no es cierto? Eligiendo esta opción evitamos el peor de todos los resultados posibles, e incluso puedo quedar libre. Naturalmente, el otro prisionero llegará a la misma conclusión que yo, así que él, probablemente, jugará sobre seguro y me traicionará, por lo cual yo debo traicionar el compromiso para evitar un desastre —z menos que esté tan loco que no me importe pasar toda la vida en la cárcel con tal de mantener una promesa—, de modo que ambos conseguiremos una sentencia de prisión de duración media. ¡Qué bien me iría todo si sólo fuera yo el que supiera desarrollar este razonamiento y cooperara!

Lo que cuenta es la estructura lógica del juego, no su anécdota particular que nos ayuda gráficamente a orientar la imaginación. Podemos sustituir las sentencias por resultados positivos (como una oportunidad de ganar diferentes sumas de dinero o, por ejemplo, *descendientes* en la evolución) en tanto en cuanto que los resultados positivos o negativos sean simétricos y ordenados

de tal forma que la traición de uno solo produzca más beneficios que la cooperación mutua, la cual «paga más» a cada uno que la traición mutua, la cual, a su vez, «paga más» que el beneficio que consigue el incauto cuando el otro es el único traidor. (Y en una escenificación formal debemos imponer otra condición: el término medio del «pago» del incauto y de la traición mutua no debe ser mayor que el «pago» de la cooperación mutua). En cualquier lugar del mundo en el que este argumento sea puesto en escena, se está representado el *Dilema del Prisionero*.

Las exploraciones de la teoría de juegos se han realizado en muchos campos, desde la filosofía y la psicología hasta la economía y la biología. De estas aplicaciones de la teoría de los juegos a la teoría de la evolución, la de mayor influencia es la concepción de Maynard Smith de una *estrategia evolutivamente estable* (ESS, por Evolutionary Stable Strategy), una estrategia que puede no ser la *mejor* desde un punto de vista olímpico (o ¡fujiano!) pero que es no mejorable y no destruible dadas las circunstancias. Maynard Smith (1988:caps. 21 y 22) escribió una excelente introducción acerca de la teoría de los juegos en la evolución. La edición revisada del libro de Dawkins *El gen egoísta* (1989) contiene una exposición del desarrollo del pensamiento ESS en biología más o menos durante la última década, particularmente buena, cuando las simulaciones a gran escala, en ordenador, de varios modelos de la teoría de los juegos revelaron complicaciones que habían sido pasadas por alto en las primeras versiones menos realistas.

Me gustaría expresar ahora la idea esencial de una *estrategia evolucionaría estable* (ESS) de una forma más escueta. Una ESS es una estrategia que funciona bien contra las copias de uno mismo. El fundamento racional de esta estrategia es el siguiente: una estrategia con éxito es aquella que domina la población. En consecuencia, tenderá a encontrar copias de sí misma. Por lo tanto, no se establecerá con éxito a menos que funcione bien contra las copias de sí misma. Esta definición no es matemáticamente tan precisa como la de Maynard Smith, y no puede sustituirla porque es realmente incompleta. Pero tiene la virtud de englobar, intuitivamente, la idea básica de la ESS (Dawkins 1989a).

No hay duda de que los análisis de la teoría de los juegos funcionan en la teoría de la evolución. ¿Por qué, por ejemplo, son tan altos los árboles de los bosques? ¡Por la misma razón que enormes conjuntos de anuncios luminosos compiten por llamar nuestra atención a lo largo de las vías comerciales en cualquier región del país! Cada árbol mira por sí mismo y trata de conseguir tanta luz solar como sea posible.

¡Ojalá las secoyas pudieran reunirse en asamblea y ponerse de acuerdo sobre algunas restricciones razonables en su entorno y dejar de competir por la luz de sol! ¡Podrían ahorrarse la dificultad de tener que construir esos ridículos y costosos troncos, y limitarse a ser arbustos bajos y económicos, con lo cual conseguirían tanta luz solar como antes! (Dennett 1990b: 132).

Pero no pueden reunirse en asamblea; bajo esas circunstancias, la ruptura o traición de cualquier «acuerdo» cooperativo ¡ría ligada a un resultado negativo o positivo cada vez que ocurriese, de modo que los árboles se verían afectados por lo que se ha llamado la «tragedia de los bienes comunitarios» (Hardin 1968) si no existiera un aporte de luz solar esencialmente inacabable. La «tragedia de los bienes comunitarios» tiene lugar cuando hay un recurso finito, «público» o compartido, de modo que los individuos están tentados, egoístamente, a coger o disponer de más de lo que honradamente les correspondería, como sucede con los peces comestibles en los océanos. A menos que puedan alcanzarse acuerdos muy específicos y de seguro cumplimiento, el resultado tenderá a ser la destrucción del recurso. Muchas especies, desde diversos puntos de vista, se enfrentan a varias versiones del *Dilema del Prisionero*. Y nosotros, los seres humanos, nos enfrentamos, consciente o inconscientemente, a estos dilemas, a veces de forma que no podíamos haber imaginado sin la ayuda del pensamiento adaptacionista.

El *Homo sapiens* no está exento de ese tipo de conflicto genético que David Haig postulaba para explicar la huella o impresión genómica: en un nuevo e importante artículo, Haig [1993] analiza una variedad de conflictos existentes entre los genes de una mujer gestante y los de su embrión. Naturalmente, conviene al embrión que la madre se mantenga fuerte y saludable, porque su propia supervivencia depende no sólo de que ella lleve a término su embarazo sino de que se haga cargo del recién nacido. Sin embargo, si la madre, en su intento de mantenerse saludable bajo circunstancias difíciles —el hambre, por ejemplo, lo que debe haber sido una circunstancia frecuente en la mayoría de las generaciones humanas—, pudiera interrumpir la nutrición que provee a su embrión, esto llegaría a ser una amenaza mayor para la supervivencia del embrión que la otra alternativa (una madre debilitada).

Si al embrión «le dieran a escoger» entre, por un lado, ser abortado de forma espontánea o nacer muerto o con poco peso, y, por otro, nacer con peso normal de una madre debilitada o incluso moribunda, ¿qué le dictaría (egoístamente) su razón? Le dictaría dar todos los pasos necesarios para tratar de salir lo menos malparado posible de la situación, evitaría que la madre cortara por lo sano (ella siempre tiene la posibilidad de tener otro hijo más tarde, cuando haya desaparecido el hambre) y esto es lo que el embrión hace. Tanto el embrión como la madre pueden quedar tan completamente al margen de este conflicto como los árboles que crecen competitivamente en el bosque. El conflicto se establece en los genes y en su control de las hormonas, no en

el cerebro de la madre ni en el del embrión; es el mismo tipo de conflicto que hemos visto entre los genes maternos y paternos en el ratón. Hay un flujo de hormonas; el embrión produce una hormona que estimula su propio crecimiento a expensas de las necesidades nutritivas de la madre; el cuerpo de la madre responde con una hormona antagonista que intenta deshacer el efecto de la primera hormona; y así sucesivamente en una escalada que puede producir niveles hormonales mucho más elevados de lo normal. Este tira y afloja habitualmente termina en un *statu quo* mutuamente satisfactorio a medias, pero origina una serie de productos colaterales que serían absolutamente desconcertantes y carentes de sentido si no fuera porque se trata de los efectos predecibles de tal conflicto. Haig concluye con una aplicación de la tesis fundamental de la teoría de los juegos: «Los genes maternos y fetales se beneficiarían si se consiguiera una transferencia de recursos con una menor producción de… hormonas y menor resistencia maternal, pero tal acuerdo no se puede forzar desde el punto de vista evolutivo» (Haig 1993:518).

Estas no son, desde muchos puntos de vista, noticias bienvenidas. Los casuales comentarios de Von Neumann acerca de la inevitabilidad del egoísmo humano resumen el punto de vista darwiniano que mucha gente ve con repugnancia sin que resulte difícil saber por qué. Temen que la expresión darwiniana «supervivencia del mejor adaptado» pueda *acarrear* que la gente sea maliciosa y egoísta. ¿Es esto justamente lo que está diciendo Von Neumann? No, de ningún modo. Está diciendo algo que está naturalmente vinculado al darwinismo, que en *general* tales virtudes como la cooperación «no se pueden cumplir desde el punto de vista evolutivo» y, en consecuencia, son difíciles de alcanzar. Si la cooperación y las otras virtudes no egoístas existen, *deben ser diseñadas*, ya que no se consiguen gratuitamente. Deben ser diseñadas bajo circunstancias especiales. (Véanse, por ejemplo, Esthel 1984 y 1985, y Haig y Graefen 1991). Después de todo, la revolución eucariota que hizo posible los organismos multicelulares fue una revolución que comenzó cuando se alcanzó una tregua entre ciertas células procariotas y ciertas bacterias invasoras. Hallaron un modo de unir fuerzas, dejando de lado sus intereses egoístas.

La cooperación y las otras virtudes son, en general, propiedades raras y especiales que solamente pueden aparecer en circunstancias de I+D muy particulares y complejas. Podemos contrastar el paradigma panglossiano con el paradigma de Pollyanna, el cual asume de manera excesivamente optimista, como hacía Pollyanna —la heroína infantil de Elisabeth Porter—, que la

madre naturaleza es hermosa[97]. En general, no lo es, aunque eso no signifique el fin del mundo. Aun en este caso, vemos que existen otras perspectivas. ¿No deberíamos considerarnos realmente afortunados, por ejemplo, de que los árboles sean tan insuperablemente egoístas? Los bellos bosques —para no hablar de los bellos barcos de vela de madera y del limpio y blanco papel en el cual escribimos nuestra poesía— no existirían si los árboles no fueran egoístas.

No cabe duda de que los diversos análisis de la teoría de los juegos funcionan en la teoría evolucionista, pero ¿funcionan *siempre*? ¿Bajo qué condiciones se aplican y cuándo podemos decir que nos estamos sobrepasando en su aplicación? Los cálculos de la teoría de los juegos siempre asumen que existe un cierto margen de movimientos posibles, a partir de los cuales los contendientes, egoístas por definición, eligen. Pero ¿hasta qué punto esto es, *en general*, realista? ¿Será porque en una circunstancia concreta es el movimiento que dicta la *razón*, es el movimiento que siempre hace la naturaleza? ¿Es éste el movimiento que siempre hará la naturaleza? ¿No es esto optimismo panglossiano? (Como ya hemos visto, esta postura parece, a veces, *pesimismo* panglossiano: «¡Qué le vamos a hacer! Los organismos son "demasiado inteligentes" para cooperar»)[98].

La hipótesis estándar de la teoría de los juegos es que siempre habrá mutaciones que poseen los efectos fenotípicos «correctos» en cada ocasión, pero ¿qué sucedería si la madre naturaleza no realiza el movimiento correcto? ¿Es esto posible de vez en cuando o a menudo? Conocemos casos en los que la madre naturaleza hace el movimiento: para hacer bosques, por ejemplo. ¿Hay, quizá, tantos casos o más en los cuales algún tipo de limitación oculta impide que esto suceda? Puede muy bien haberlos, aunque, en tal caso, los adaptacionistas desearían persistir en su idea, planteando la siguiente pregunta: ¿existe en este caso alguna razón para que la madre naturaleza no haga el movimiento, o es simplemente una limitación brutal e insensata aplicada sobre el juego racional de la madre naturaleza?

Gould ha sugerido que un defecto fundamental en el razonamiento adaptacionista es la presunción de que en todo paisaje adaptativo siempre aparece claro donde está el camino hacia las partes más elevadas de las distintas cumbres, aunque cabe suponer que haya limitaciones ocultas, como si fueran vías de ferrocarril extendidas a lo largo del paisaje. «Las limitaciones heredadas y las vías de desarrollo pueden así canalizar cualquier cambio ya que, aunque la selección induce a movimientos descendentes a través de los caminos permitidos, ese canal, en sí mismo, representa el

determinante primario de la dirección evolutiva» (Gould 1982a:383). Las poblaciones, por lo tanto, no consiguen la expansión *ad lib* en toda la amplitud del terreno, sino que son forzadas a seguir los caminos, como se representa en la figura 9.4.

Supongamos que esto sea verdad. Entonces ¿cómo se pueden localizar las limitaciones ocultas? Está muy bien que Gould y Lewontin señalen la posibilidad de que existan limitaciones ocultas —todo adaptacionista reconoce que esta es una posibilidad omnipresente—, pero es necesario considerar cuál es la mejor metodología para descubrirlas. Consideremos ahora una curiosa variación en la práctica estándar del ajedrez.

Cuando un buen jugador juega con un contrincante más débil en partidas amistosas, el jugador más fuerte se ofrece a menudo voluntariamente para jugar con una desventaja —un *handicap*—, para que la partida se iguale y sea más excitante. El *handicap* estándar es darle una pieza o dos de ventaja; jugar solamente con un alfil o una torre o, en casos extremos, sin la reina. Pero existe otro sistema de *handicap* que puede dar resultados interesantes. Antes de la partida, el jugador más fuerte anota en un trozo de papel una limitación oculta (o varias limitaciones) bajo la cual ella se comprometerá a jugar, y esconderá el papel bajo el tablero. ¿Cuál es la diferencia entre una limitación y un movimiento forzado? La razón dicta un movimiento forzado —y siempre lo dictará una y otra vez— mientras que algún congelado *bit* de la historia dictará una limitación, hubiera o no una razón para su nacimiento y haya o no ahora una razón a favor o en contra. He aquí una serie de limitaciones posibles:

A menos que me vea forzado por las reglas para hacerlo así (debido a que estoy en jaque y por tanto obligado a realizar cualquier movimiento legal para escapar del mismo).

1) No debo nunca mover la misma pieza dos veces consecutivas.
2) No me puedo enrocar.
3) Sólo puedo capturar piezas con peones tres veces durante toda la partida.
4) Mi reina solamente se puede mover como una torre y nunca en diagonal.

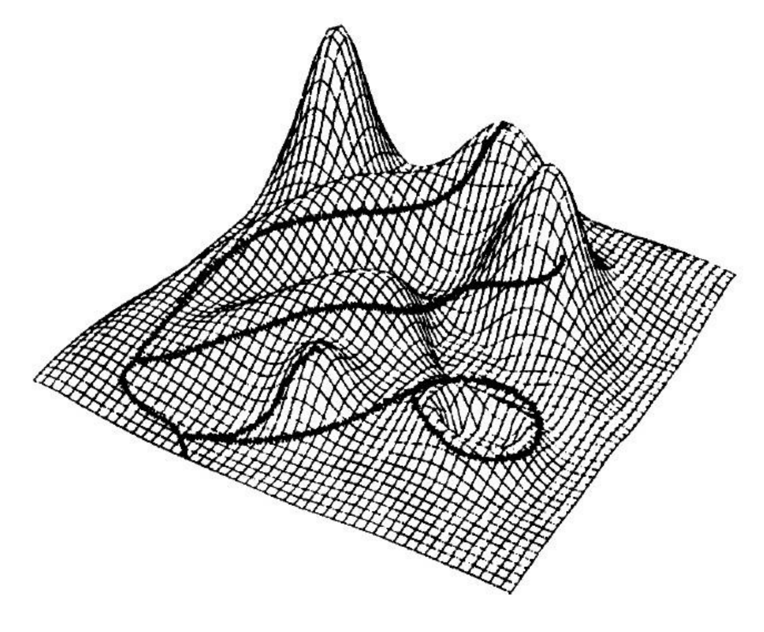


Figura 9.4

Imaginemos ahora el dilema epistemológico del jugador más débil, que sabe que su oponente está jugando con limitaciones ocultas aunque no *sabe* cuáles son. ¿Cómo debe proceder? La respuesta es totalmente obvia: debe jugar *como* si todos los movimientos *aparentemente* posibles —todos los movimientos legales— estén disponibles para ella —el jugador más fuerte— y ajustar su estrategia solamente cuando comience a aumentar la evidencia de que le está realmente prohibido hacer lo que de otra manera sería obviamente el mejor movimiento.

Esta evidencia no es, en ningún caso, fácil de reunir. Si pensamos que nuestra oponente no puede mover su reina en diagonal, cabe probar esta hipótesis mediante la arriesgada táctica de ofrecer una pieza gratis a la reina precisamente en la diagonal. Si la reina declina el ofrecimiento, esto cuenta a favor de nuestra hipótesis; *a menos* que haya una razón más profunda de estrategia (inimaginable hasta ahora para nosotros) para declinar la captura de la pieza regalada. (Recordemos la segunda regla de Orgel: la evolución es más lista que nosotros).

Naturalmente, otro modo de conocer cuáles son las limitaciones ocultas en el tablero de ajedrez es mirar el papel escondido; y uno puede pensar que lo que Gould y Lewontin están recomendando es que los adaptacionistas simplemente abandonen el juego que están jugando y alcancen la verdad a través del examen más directo de la evidencia molecular. Desafortunadamente, esta analogía es engañosa. Estamos ciertamente

capacitados para utilizar cualquier habilidad para reunir datos disponibles en el juego de la ciencia, pero cuando miramos de reojo a las moléculas, todo lo que encontramos es más maquinaria, más diseño (o aparente diseño), que precisa de la aplicación de la ingeniería revertida. En ninguna parte están escritas las limitaciones ocultas de la madre naturaleza de tal forma que puedan ser leídas sin la ayuda de reglas interpretativas de la hermenéutica de los artefactos (Dennett 1990b). El descenso hasta los más profundos niveles del ADN, por ejemplo, es naturalmente un método valioso para mejorar ampliamente la precisión de la investigación —aunque habitualmente con el intolerable coste de anegarse en datos excesivos— pero, en cualquier caso, no es ésta una alternativa para el adaptacionismo, sino una extensión de éste.

El ejemplo de jugar al ajedrez con limitaciones ocultas nos permite comprobar que existe una profunda diferencia entre la madre naturaleza y los jugadores humanos de ajedrez, la cual creo que tiene implicaciones para una extendida debilidad del pensamiento adaptacionista. Si *estamos* jugando al ajedrez bajo limitaciones ocultas deberemos adaptar nuestra estrategia a esta situación. Sabiendo que hemos prometido secretamente no mover la reina en diagonal, olvidaremos todo plan que coloque a la reina en situación de riesgo de captura gracias a sus inusuales limitaciones; aunque naturalmente podemos correr riesgos, esperando que nuestro débil oponente no tenga noticia de esta posibilidad. Pero nosotros sí conocemos la existencia de limitaciones ocultas y las tenemos en cuenta. La madre naturaleza no. La madre naturaleza no tiene razón para evitar maniobras de elevado riesgo; ella asume todos los riesgos y cuando pierde, en la mayoría de los casos, se encoge de hombros.

Aquí se aplica esta idea en el pensamiento evolucionista. Supongamos que tenemos noticia de que un tipo particular de mariposa posee en sus alas una coloración que milagrosamente mimetiza las combinaciones de colores del suelo del bosque donde vive. Atribuimos este hecho a una hermosa adaptación, a un camuflaje, lo que sin duda es. Esta mariposa lo hace mejor que sus primas *debido a que* su coloración reproduce perfectamente la del suelo del bosque. Pero hay una tentación a la cual se sucumbe, la de añadir, implícita o explícitamente: «¡Y lo que es más, si el suelo del bosque tuviera otra coloración, la mariposa tendría esa otra combinación de colores!». Esta afirmación es impertinente. Puede muy bien no ser verdad. En última instancia, podría ser que esta coloración del suelo del bosque fuera el *único* tipo de coloración que este linaje de mariposa puede mimetizar con suficiente éxito; si el suelo del bosque fuera diferente, ese linaje no estaría allí (no olvidar nunca en la evolución la importancia del «palo y la zanahoria»). Si el

suelo del bosque cambia, ¿qué sucederá? ¿Se adaptaría automáticamente la mariposa? ¡Lo más que podemos decir es que la mariposa se adaptará cambiando su camuflaje o no! Si no cambia podría encontrar alguna otra adaptación en su limitado arsenal de movimientos disponibles, o de lo contrario desaparecería muy pronto.

El caso límite, en el cual sólo un camino está abierto a la exploración, es un ejemplo del actualismo, nuestra vieja némesis: sólo lo actual es posible. Tales exploraciones encorsetadas del espacio de la (aparente) posibilidad no están descartadas, en mi opinión, pero deben ser la excepción, no la regla. Si hubieran sido la regla, el darwinismo habría muerto, absolutamente incapaz de explicar nada (aparente) del diseño en la biosfera. ¡Sería como si hubiéramos escrito un programa para jugar al ajedrez en el ordenador que pudiera jugar una partida por repetición mecánica (por ejemplo, los movimientos de Alekhine en la famosa partida Flamberg-Alekhine, jugada en Manheim en 1914) y, *mirabile dictu*, que le ganase sistemáticamente a todos sus adversarios! Esto sería una «armonía preestablecida» de proporciones milagrosas, algo así como una burla de la tesis de Darwin que afirma tener una explicación de cómo han sido encontrados los movimientos «ganadores».

Pero nuestro abandono del actualismo no debe hacernos caer en el error opuesto, el de suponer que el espacio de las posibilidades reales se encuentra más densamente poblado de lo que realmente está. La tentación, cuando pensamos acerca de variación fenotípica, es adoptar una suerte de táctica similar a la aplicada en el retrato robot de sospechosos (procedimiento «Identikit»), en el cual se asume que todas las variaciones menores que podamos imaginar sobre los temas que encontramos en la realidad están verdaderamente disponibles. Llevada a estos extremos, esta táctica estaría siempre sobreestimando inmensamente —«vastamente»— lo que es realmente posible. Si el árbol de la vida *actual* ocupa evanescentes estrechas hileras a través de la Biblioteca de Mendel, el árbol de la vida *realmente posible* es más bien un frondoso arbusto pero aún lejos de los árboles densamente frondosos de lo *aparentemente posible*. Hemos visto ya que el «vasto» espacio de todos los imaginables fenotipos —podemos llamarlo «Identikit Space»— incluye sin duda grandes regiones para las cuales no hay recetas en la Biblioteca de Mendel. Pero, incluso a lo largo de los caminos a través de los cuales el árbol de la vida se extiende, no tenemos garantías de que las regiones vecinas del «Identikit Space» sean realmente accesibles en su totalidad**[99]**.

Si las limitaciones ocultas garantizan que haya un gran conjunto de muros de un laberinto, en gran parte invisible —o canales o vías de ferrocarril—, en el espacio de las posibilidades aparentes, entonces la expresión «no podemos llegar allí desde aquí» será verdadera mucho más a menudo de lo que podemos imaginar. Incluso, si esto es así, no podemos hacer nada mejor en nuestra exploración de esta posibilidad que poner en juego nuestras estrategias de ingeniería revertida, a cada momento y en todo nivel. Es importante no sobreestimar las posibilidades reales, pero todavía es más importante no subestimarlas, un punto débil frecuente, aunque en ninguno de los adaptacionistas se manifiesta habitualmente. Muchos argumentos adaptacionistas son del tipo «si es posible esta variedad sucederá»: los tramposos emergerán para invadir a los santos; o una «carrera de armamento» seguirá hasta que se consiga una estabilidad adaptativa de primer orden, etc. Estos argumentos presuponen que una parte suficiente del espacio de las posibilidades es «habitable» para asegurar que el proceso se aproxima al modelo de la teoría de los juegos utilizado. Pero ¿son apropiadas estas hipótesis? ¿Mutarán estas bacterias en una forma que sea resistente a nuestra nueva vacuna? No, si somos afortunados, pero será mejor asumir lo peor; es decir, que hay en el espacio actualmente accesible a estas bacterias contramovimientos en la «carrera de armamento» que nuestra innovación médica ha puesto en movimiento (Williams y Nesse 1991).

CAPÍTULO 10
¡Bravo por los brontosaurios!

I. *El niño que gritó que viene el lobo*

> Los científicos tienen poder gracias al respeto que emanan sus disciplinas. En consecuencia, podemos sentirnos penosamente tentados a utilizar indebidamente este poder en fomento de un prejuicio personal o de un objetivo social; ¿por qué no disponer de esta energía extra para extender el paraguas de la ciencia sobre una preferencia personal ética o política? Pero no podemos hacerlo, pues podríamos perder el verdadero respeto que en un principio fue la causa de la tentación.
>
> Stephen Jay Gould, *Brontosaurus y la nalga del ministro*

Hace muchos años, vi un programa de la televisión británica en el cual entrevistaban a unos niños acerca de la reina Isabel II. Sus confiadas respuestas eran encantadoras: la reina, según ellos, pasaba gran parte del día pasando la aspiradora por el palacio de Buckingham, con la corona puesta, naturalmente. Acercaba su trono a la televisión cuando no estaba ocupada en asuntos de Estado y utilizaba un delantal sobre los trajes de armiño cuando hacía la limpieza. Llegué a la conclusión de que la muy imaginaria reina Isabel II de estos niños (lo que los filósofos llamarían su *objeto intencional*) era, en algunos aspectos, un objeto más potente e interesante en el mundo que la mujer real. Los *objetos intencionales* son criaturas de creencias, y en consecuencia, desempeñan un papel más directo en guiar (o en guiar erróneamente) la conducta de la gente que el que desempeñan los objetos reales que aparentan ser idénticos. El oro de Fort Knox, por ejemplo, es menos importante de lo que se cree y el mito de Albert Einstein es, como Santa Claus, bastante mejor conocido que el personaje histórico recordado de manera vaga que fue la fuente primaria de este mito.

Este capítulo trata de otro mito, Stephen Jay Gould, el refutador del darwinismo ortodoxo. A lo largo de los años, Gould ha puesto en marcha una serie de ataques contra diversos aspectos del neodarwinismo contemporáneo y aunque ninguno de estos ataques ha probado ser a lo sumo algo más que una leve corrección de la ortodoxia, su impacto retórico sobre el mundo exterior

ha sido inmenso y distorsionante. Este hecho me plantea un problema que no puedo pasar por alto ni posponer. En mi propio trabajo, a lo largo de los años, he apelado a menudo a consideraciones evolutivas y he corrido, casi con la misma frecuencia, contra una corriente de resistencia: mis apelaciones a favor del razonamiento darwiniano han sido rudamente rechazadas como ciencia desacreditada y pasada de moda, por filósofos, psicólogos, lingüistas, antropólogos y otros, quienes me han hecho saber, jocosamente, que toda mi biología era errónea; yo no había hecho mis deberes, puesto que Steve Gould había demostrado que el darwinismo, después de todo, no se encontraba en tan buena forma, sino, por el contrario, muy cerca de la extinción.

Esto es un mito, pero un mito con mucha influencia, incluso en los templos de la ciencia. En este libro, he tratado de presentar una visión ajustada del pensamiento evolucionista, apartando al lector de los malentendidos más frecuentes, y defendiendo la teoría contra las objeciones mal fundamentadas. He contado con mucha ayuda y consejos de expertos y, por ello, confío en tener éxito. Pero el punto de vista del pensamiento darwiniano que he presentado está completamente en desacuerdo con el punto de vista convertido por Gould en familiar para muchos. Entonces, ¿está equivocado mi punto de vista? Después de todo, ¿quién conoce mejor que Gould a Darwin y al darwinismo?

Los norteamericanos están notoriamente mal informados acerca de la evolución. Una encuesta Gallup (junio de 1993) ha detectado que el 47 por ciento de los norteamericanos adultos cree que el *Homo sapiens* es una especie creada por Dios hace menos de diez mil años. De todos modos, que los norteamericanos conozcan algo sobre la evolución se debe, probablemente, más a Gould que a ningún otro. En la batalla por enseñar la «ciencia de la creación» en las escuelas, ha sido un testigo clave en la defensa de la evolución en los tribunales de justicia, una persistente plaga en la educación norteamericana. Durante veinte años, la columna mensual de Gould titulada «La visión de la vida» en *Natural History* ha aportado a los biólogos aficionados y profesionales una continua corriente de puntos de vista cautivadores, de hechos fascinantes y de correcciones muy necesarias para sus pensamientos. Además de esta colección de ensayos, recogidos en volúmenes tales como *Desde Darwin* (1977), *El pulgar del panda* (1980), *Dientes de gallina, dedos de caballo* (1983), *La sonrisa del flamenco* (1985), *Brontosaurus y la nalga del ministro* (1991) y *Ocho cerditos* (1993) y de sus publicaciones técnicas sobre caracoles y paleontología, Gould ha escrito un gran libro teórico, *Ontogeny and Phytogeny* (1977); un ataque sobre las

pruebas para la medida de la inteligencia, *La falsa medida del hombre* (1981); un libro sobre la reinterpretación de la fauna encontrada en el yacimiento de fósiles del Burgess Shale, *La vida maravillosa* (1989); y numerosos artículos sobre temas que abarcan desde Bach hasta el béisbol, desde la naturaleza del tiempo a los compromisos de *Parque jurásico*. La mayoría de estas obras son sencillamente admirables: sorprendentemente eruditas, Gould aparece en ellas como el verdadero modelo de un científico que reconoce, como decía mi profesor de física en la enseñanza secundaria, que la ciencia, bien hecha, es una de las humanidades.

En realidad, el título de la columna mensual de Gould procede de Darwin; es la frase final de *El origen de las especies*.

Hay grandeza en esta visión de que la vida, con sus diferentes poderes, ha sido originalmente infundida en unas pocas formas o en una sola; y que mientras este planeta ha estado girando de acuerdo con la constante ley de la gravedad, a partir de un comienzo tan simple, infinidad de formas más bellas y más maravillosas han evolucionado y siguen evolucionando.

Alguien tan prolífico y activo como Gould tendrá seguramente una agenda de trabajo que va más allá de educar y deleitar simplemente a los seres humanos con la visión darwiniana de la vida. De hecho, Gould ha tenido numerosas agendas. Ha luchado con dureza contra el prejuicio y, especialmente, contra el abuso de la investigación científica (y del prestigio científico) a causa de aquellos que cubren sus ideologías políticas bajo el potente manto de la respetabilidad de la ciencia. Es importante reconocer que el darwinismo ha tenido siempre un infortunado poder de atracción sobre muchos entusiastas que no son bienvenidos: demagogos, psicópatas, misántropos y otros que abusan de la peligrosa idea de Darwin. Gould ha puesto al descubierto esta triste situación en docenas de relatos sobre el darwinismo social, los racistas incalificables y, más particularmente sobre la gente, buena en el fondo, que cae en la confusión, seducida y abandonada por una u otra sirena darwiniana. Como quiera que es demasiado fácil huir precipitadamente con una versión pobremente comprendida del pensamiento darwiniano, Gould ha cumplido, durante la mayor parte de su vida, el trabajo de proteger a su héroe de este tipo de abuso.

La ironía es que a Gould, con sus persistentes esfuerzos por proteger al darwinismo, a veces le ha salido el tiro por la culata. Gould ha sido un defensor de su propia rama del darwinismo, pero un ardiente opositor de lo que él ha llamado «ultradarwinismo» o «hiperdarwinismo». ¿Cuál es la diferencia? El darwinismo no comprometido que he presentado, aquel que «no admite ganchos en el cielo», a los ojos de Gould es «hiperdarwinismo»,

es decir, una visión extremista que, para él, es necesario rechazar. Dado que de hecho se trata, como ya he dicho, de neodarwinismo completamente ortodoxo, las campañas de Gould han tomado la forma de llamadas a la revolución. Desde su intimidante púlpito, Gould ha anunciado, una y otra vez, a un mundo de fascinados espectadores, que el neo-darwinismo ha muerto para ser suplantado por una nueva visión revolucionaria —aún darwiniana— pero que rechaza el punto de vista del *establishment*. Sin embargo esta revolución anunciada no ha tenido lugar. Como dice Simon Conway Morris, uno de los héroes del libro de Gould titulado *La vida maravillosa*: «Sus puntos de vista han hecho todo lo posible por agitar las ortodoxias establecidas, pero aun así, cuando la polvareda se asienta, el edificio de la teoría de la evolución permanece casi sin cambios» (Conway Morris 1991:6).

Gould no es el único evolucionista que ha sucumbido al impulso de una dramatización excesiva. Manfred Eigen y Stuart Kauffman —y otros que no hemos considerado— en un principio se presentaron a sí mismos como herejes radicales. ¿Quién no preferiría una contribución verdaderamente revolucionaria? Pero mientras que Eigen y Kauffman, como hemos visto, han moderado sus teorías a su debido tiempo, Gould ha ido de revolución en revolución. A la larga, sus llamadas a la revolución han sido todas falsas alarmas, aunque él ha continuado clamando, en franco desafío a la moraleja de la fábula de Esopo sobre el niño que gritaba que venía el lobo. Esta actitud no sólo le ha creado un problema de credibilidad (entre los científicos), sino que también ha concitado la animosidad de algunos de sus colegas, quienes han sentido el aguijón de lo que ellos consideran una inmerecida condenación pública provocada por sus influyentes campañas. Como puntualiza Robert Wright, Gould es «el laureado evolucionista norteamericano. Que haya estado confundiendo sistemáticamente a los norteamericanos acerca de lo que es la evolución y lo que significa, ha supuesto un enorme daño intelectual» (Wright 1990:30).

Realmente, ¿ha hecho esto Gould? Veamos lo que sigue. Si creemos que 1) el adaptacionismo ha sido refutado o relegado a un papel menor en la biología evolucionista, o 2) dado que el adaptacionismo es «el fallo intelectual fundamental de la sociobiología» (Gould 1993a:319) la sociología ha sido absolutamente desacreditada como disciplina científica, o 3) la hipótesis de Gould y Eldredge del «equilibrio puntuado» descarta el neodarwinismo ortodoxo, o 4) Gould ha demostrado que el hecho de la extinción en masa refuta el «extrapolacionismo» que es el talón de Aquiles del neo-darwinismo ortodoxo, entonces, *lo que creemos es una falsedad*. Sin

embargo, si creemos cualquiera de estas proposiciones, estamos en muy buena compañía, una numerosa e intelectualmente distinguida compañía. En cierta ocasión Quine dijo de un crítico que había leído mal su obra: «Lee con brocha gorda». Todos somos capaces de hacerlo así, sobre todo cuando tratamos de construir, para nuestro propio uso, en términos simples, el mensaje, de trabajos que caen fuera de nuestro propio campo. Tenemos tendencia a leer, con trazos gruesos, lo que deseamos encontrar. Cada una de estas cuatro proposiciones expresa un veredicto que es más decisivo y radical de lo que Gould intentaba, pero juntas constituyen un mensaje que se ha extendido en muchos campos. Estoy en desacuerdo, así que me toca desmantelar el mito. No es un trabajo fácil, dado que debo separar claramente la forma retórica de la realidad, mientras que defiendo —explicándolo— la presunción, totalmente razonable, de que un evolucionista de la categoría de Gould no puede estar equivocado en sus veredictos, o ¿acaso puede estarlo? Sí y no. El verdadero Gould ha hecho grandes contribuciones al pensamiento evolucionista, corrigiendo una variedad de graves interpretaciones ampliamente extendidas, pero el mítico Gould ha sido creado por los que le temen a Darwin, alimentados por las muy apasionadas palabras de Gould, y esto ha estimulado, a su vez, sus propias aspiraciones de destruir el «ultradarwinismo», lo que le ha conducido a afirmaciones espurias.

Si Gould ha seguido gritando que viene el lobo, ¿por qué lo ha hecho? La hipótesis que defiendo es que Gould está siguiendo una larga tradición de eminentes pensadores que han estado buscando «ganchos celestes» pero se encuentran con grúas. Dado que la teoría de la evolución ha hecho grandes progresos en años recientes, la tarea de buscar un hueco para los «ganchos celestes» se ha hecho más difícil, subiendo el listón para cualquier pensador que desee encontrar alguna bendita excepción. El seguimiento, en el trabajo de Gould, de la repetición del tema y sus variaciones descubre una pauta de actuación: cada intento fallido define una pequeña porción de la sombra de su presa, hasta que eventualmente queda claramente dibujada la fuente de la incomodidad que estimula a Gould. El último objetivo de Gould es la peligrosa idea de Darwin en sí misma. Gould se opone a la simple idea de que la evolución sea, en último término, un proceso algorítmico.

Sería interesante preguntar, por último, por qué Gould está tan firmemente asentado en su oposición a la idea de Darwin, pero ésta es una tarea que dejo para otra ocasión y, quizá, para otro escritor. El propio Gould ha mostrado cómo llevar a cabo esta tarea: ha examinado las hipótesis subyacentes, los miedos y las esperanzas de los primeros científicos, desde el mismo Darwin

pasando por Alfred Binet, el inventor del *test* IQ, para la medida del coeficiente intelectual, hasta Charles Walcott, el clasificador (erróneo) de la fauna hallada en el yacimiento de fósiles de Burgess Shale, para citar tres de las historias mejor conocidas. ¿Qué planteamientos ocultos —morales, políticos, religiosos— han guiado al propio Gould? Aunque esta es una pregunta fascinante, voy a resistir la tentación de responderla, sin perjuicio de que, a su debido tiempo, considere brevemente, como es debido, las hipótesis rivales que se han sido sugerido. Tengo bastante faena con la defensa de mi declaración inicial de que el modelo de las revoluciones fallidas de Gould hace patente que el laureado evolucionista norteamericano se ha encontrado siempre incómodo con el núcleo fundamental del darwinismo.

Durante años me he sentido desconcertado ante la mal definida hostilidad hacia el darwinismo que he encontrado en muchos de mis colegas académicos, quienes mientras citaban a Gould como autoridad, a mi parecer lo leían de una manera equivocadamente voluntarista, con una pequeña ayuda de los medios de comunicación, siempre dispuestos a destruir con sutileza y a avivar las llamas de cualquier controversia menor. Esto no me ocurrió realmente a mí, ya que Gould estaba, a menudo, luchando en el otro campo. El mismo Gould ha sido victimizado, con frecuencia, debido a esta hostilidad. Maynard Smith menciona precisamente un ejemplo:

Uno no se puede pasar la vida trabajando en la teoría de la evolución sin llegar a darse cuenta de que la mayoría de la gente que no trabaja en este campo, y alguna que sí lo hace, tiene una fuerte tendencia a aceptar que la teoría darwiniana es falsa. Esta consideración me vino a la mano cuando mi amigo Stephen Jay Gould, que es un darwinista tan convencido como yo, aprovechó la oportunidad de un editorial en el *Guardian* para anunciar la muerte del darwinismo, seguido de una extensa correspondencia sobre el mismo tema, simplemente porque había hallado algunas dificultades, a las cuales la teoría aún se enfrenta. (Maynard Smith 1981:221; reimp., 1988).

¿Por qué un «darwinista tan convencido» como Gould sigue complicándose la vida al contribuir a la falsa concepción, extendida en la opinión pública, de que el darwinismo ha muerto? No hay un adaptacionista más comprometido y brillante que John Maynard Smith, pero aquí, creo, hemos cogido al maestro «echando una siesta»: no se interroga a sí mismo sobre la cuestión del «porqué». Después de que yo comenzara a darme cuenta de que muchas de las más importantes contribuciones a la teoría evolucionista se debían a pensadores que se encontraban incómodos con la gran idea de Darwin, pude empezar a tomar en serio la hipótesis de que el propio Gould era uno de ellos. Desarrollar esta hipótesis llevará trabajo y paciencia, pero no vamos a evitarlo. La mitología alrededor de lo que Gould ha demostrado y no ha

demostrado está tan extendida que oscurecerá todos los demás temas a los que nos enfrentamos, si antes no hacemos todo lo posible por dispersarla.

2. *Las pechinas de San Marcos y el pulgar del panda*

> Pienso que se puede comprobar lo que se está resquebrajando en la teoría de la evolución: la estricta construcción de la moderna síntesis con sus creencias en la adaptación penetrante, el gradualismo y la extrapolación por suave continuidad a causa de cambios en poblaciones locales hasta tendencias predominantes y transiciones en la historia de la vida.
>
> Stephen Jay Gould,
> *«Is a New and General Theory of Evolution Emerging?»*

> El tema no es la idea general de que la selección natural pueda actuar como una tuerza creativa; el argumento básico, en principio, está fundamentado. La duda primaria se centra en las tesis secundarias: el gradualismo y el programa adaptacionista.
>
> Stephen Jay Gould,
> *«Darwinism and the Expansion of Evolutionary Theory»*

> Gould ha trabajado bastante con el fin de colocar en primer plano un tema central del darwinismo, que la supuesta perfección en el diseño es un compromiso alcanzado con el jurado adoptando improbables piezas de anatomía. Pero algunos de estos ensayos contienen alusiones a que la explicación darwiniana es sólo parcialmente correcta. Pero ¿es este un ataque serio? No lo es, en una lectura más atenta.
>
> Simon Conway Morris,
> *«Rerunning the Tape»*

Gould (1980b, 1982a) observa dos problemas principales en la síntesis darwiniana moderna: la «adaptación invasiva» y el «gradualismo». Y los considera como problemas relacionados. ¿Cómo? Gould ha dado respuestas algo diferentes a lo largo de los años. Podemos comenzar con la «adaptación invasiva». Para conocer lo que es este tema debemos volver al trabajo de Gould y Lewontin de 1979. El título evoca un buen lugar donde comenzar: «The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme» (‘Las pechinas de San Marcos y el paradigma panglossiano: Una crítica del programa adaptacionista’). Además de redefinir el término «panglossiano», los autores introducen otro término que se ha hecho famoso: *spandrel* (pechina) ya que ha sido una acuñación de enorme éxito en un sentido: se ha extendido a lo largo de la biología evolucionista y

más allá. En un reciente ensayo retrospectivo, Gould hace referencia a esta difusión:

> Diez años más tarde, mi amigo Dave Raup me dijo: «Todos hemos sido *spandrelized*». Cuando nuestro ejemplo se convierte a la vez en genérico y en una parte diferente del discurso, hemos ganado. A partir de este momento podemos llamar a las pechinas de San Marcos «Kleenex», «Jell-O» y, con un término más enfáticamente no metafórico, «Band-Aid» (Gould 1993a:325).

Desde Gould y Lewontin, los evolucionistas (y muchos otros) han hablado de pechinas pensando que sabían de lo que estaban hablando. ¿Qué son las pechinas? Una buena pregunta. Gould desea convencernos de que la adaptación no es «invasiva», así que necesita un término para designar los hechos biológicos (presumiblemente muchos) que no son adaptaciones. Los denomina «pechinas». Las pechinas son cosas que *no son* adaptaciones, sean lo que sean. ¿Han demostrado Gould y Lewontin que las pechinas son ubicuas en la biosfera? No lo han hecho. Una vez que aclaremos las confusiones acerca de lo que significa el término utilizado Gould y Lewontin podremos ver que estos *spandrels* no son ubicuos en modo alguno o son la *base normal* para las adaptaciones y, en consecuencia, un epítome de la «adaptación invasiva».

El artículo de Gould y Lewontin comienza con dos famosos ejemplos arquitectónicos y teniendo en cuenta que al principio los autores dan un mal paso que es crucial, debemos analizar cuidadosamente el texto. (Uno de los efectos de los textos clásicos es que la gente los recuerda mal, por haberlos leído una vez y con prisas. Incluso si estamos familiarizados con el inicio del texto, debemos releerlo lentamente, para comprobar cómo se produce el mal paso, ante nuestros propios ojos).

La gran cúpula de la Catedral de San Marcos, en Venecia, presenta en sus mosaicos una detallada iconografía que expresa los pilares de la fe cristiana. Tres círculos de figuras irradian desde una imagen central de Cristo: ángeles, discípulos y virtudes. Cada círculo está dividido en cuadrantes, aun cuando la estructura de la cúpula es radialmente simétrica. Cada cuadrante se reúne con una de las cuatro pechinas en los arcos situados bajo la cúpula. Las pechinas —espacios triangulares formados por la intersección de dos arcos redondeados en ángulo recto (véase la figura 10.1)— son productos secundarios necesarios desde el punto de vista arquitectónico, para montar una cúpula sobre arcos redondos. Cada pechina contiene una ornamentación admirablemente ajustada en el espacio triangular del vértice inferior El diseño

es tan elaborado, armonioso y pleno de propósito que uno se siente tentado de verlo como el punto de inicio del análisis, como la causa, en algún sentido, de la arquitectura que le rodea. Pero esto sería invertir el orden que ha de seguir el análisis. El sistema comienza con una limitación arquitectónica: las necesarias cuatro pechinas y su forma triangular. Estas pechinas, con el vértice hacia abajo, proporcionan un espacio en el cual trabajaron los constructores de los mosaicos; las pechinas completan la simetría cuatripartita de la bóveda situada por encima.



Figura 10.1 - Una de las pechinas de la iglesia de San Marcos.

Todo techo con una bóveda en abanico debe tener una serie de espacios abiertos a lo largo de la línea media de la bóveda, donde los extremos de los abanicos se entrecruzan entre los pilares (véase la figura 10.2.). Dado que estos espacios deben existir, a menudo se utilizan para obtener ingeniosos efectos ornamentales. Por ejemplo, en la capilla del King’s College, en Cambridge, los espacios contienen tachones de adorno embellecidos alternativamente con la rosa de los Tudor y con rastrillos. De algún modo, este diseño representa una «adaptación», pero la limitación arquitectónica es evidentemente primaria. Los espacios surgen como un producto secundario de la bóveda en abanico; su uso apropiado es un efecto secundario. Cualquiera que trate de argumentar que la estructura existe debido a la alternancia de rosas y de rastrillos en una capilla Tudor quedaría tan en ridículo como hizo quedar Voltaire al doctor Pangloss… Los biólogos evolucionistas, con su tendencia a fijar su atención exclusivamente en la adaptación inmediata a las condiciones locales, tienden a pasar por alto las limitaciones arquitectónicas y realizan de este modo una inversión de la explicación (Gould 1993a:147-149).

En primer lugar, debemos señalar que, desde el comienzo, Gould y Lewontin nos invitan a *contrastar* el adaptacionismo con una preocupación por la «necesidad» arquitectónica o «limitación», como si el descubrimiento de estas limitaciones no fuera una parte integral del buen razonamiento adaptacionista, como ya he argumentado en los dos últimos capítulos. En este momento, quizá debamos detenernos y considerar la posibilidad de que Gould y Lewontin hayan sido mal interpretados por la mayoría de los que los han leído, debido a la retórica extemporánea de su pasaje inicial, retórica que corrigen, de algún modo, en la última frase antes citada. Quizá lo que Gould y Lewontin demostraron en 1979, es que todos debemos ser *mejores* adaptacionistas; debemos *extender* nuestra perspectiva de ingeniería revertida hacia atrás hasta los procesos de I+D y del desarrollo embriológico, en lugar de limitarnos «exclusivamente a la adaptación *inmediata* a las condiciones *locales*». Esta es, después de todo, una de las principales lecciones de los dos últimos capítulos, y Gould y Lewontin podrían compartir el crédito de presentar esta cuestión a la atención de los evolucionistas. Pero casi todo lo que Gould y Lewontin han dicho milita contra esta interpretación: es decir, se opone al adaptacionismo, no lo incrementa. Sus enseñanzas están a favor de un «pluralismo» en biología evolucionista, en el cual el adaptacionismo es justamente un elemento, con su influencia disminuida por otros elementos, si no absolutamente suprimida.

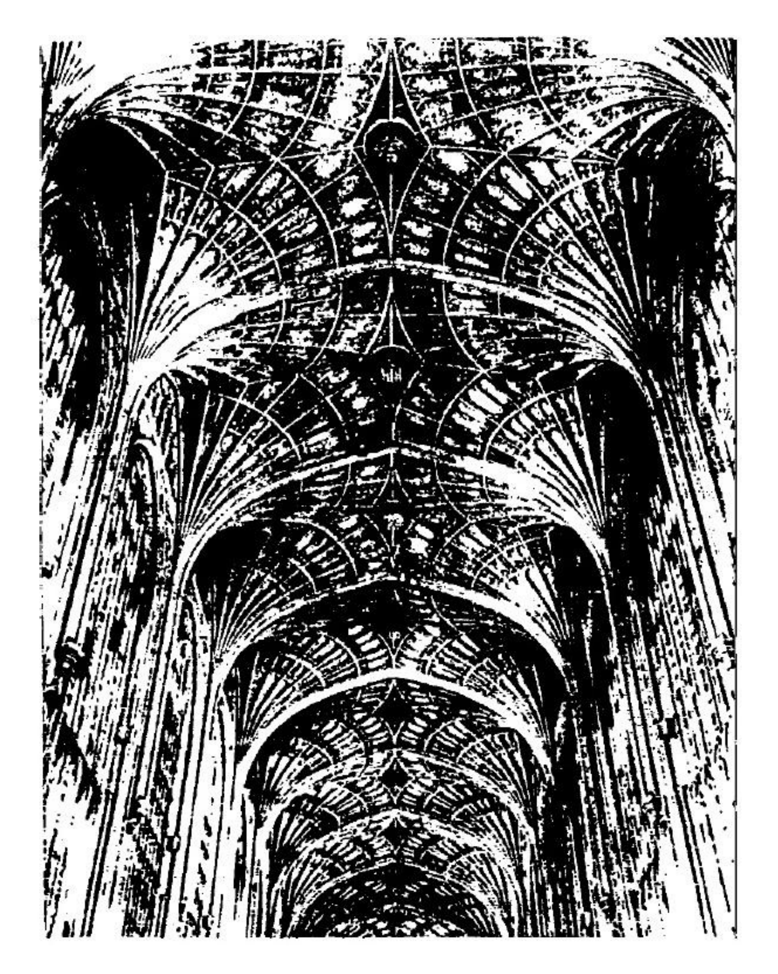


Figura 10.2 - El techo de la capilla del King’s College.

Las pechinas de San Marcos, se nos dice, «son productos arquitectónicos secundarios y necesarios para asentar una cúpula sobre arcos redondeados». ¿En qué sentido son necesarios? La suposición habitual entre los biólogos a los que he consultado es que ésta es una necesidad *geométrica* y, en consecuencia, no tiene nada que ver con los cálculos de coste-beneficio de los adaptacionistas, dado que, sencillamente, no hay elección: ¡las pechinas hay que hacerlas! Como escriben Gould y Lewontin, «las pechinas deben existir una vez que los planos de la edificación especifican que una cúpula ha de apoyarse sobre arcos redondeados» (1979:161). Pero ¿esto es así verdaderamente? A primera vista, puede parecer como si no hubiera alternativas a esas lisas superficies triangulares de vértice inferior situadas entre la cúpula y los cuatro arcos redondeados, aunque, de hecho, hay muchísimas formas de rellenar estos espacios con mampostería, todas ellas con casi idéntica firmeza estructural, y de fácil construcción. En la figura 10.3 se encuentra a la izquierda el esquema de las pechinas de San Marcos y, a la

derecha, dos variantes. Las dos variantes son, en una palabra, feas (deliberadamente las he dibujado así) pero esto no las hace *imposibles*.

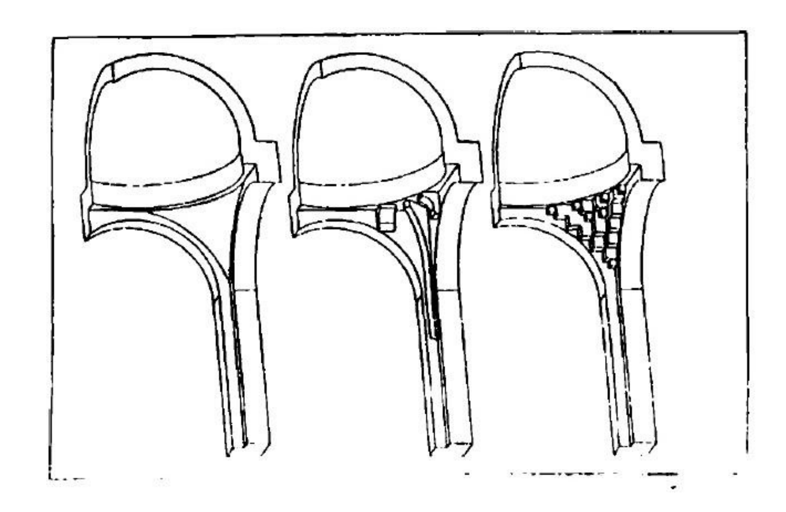


Figura 10.3

En el trabajo de Gould y Lewontin hay una confusión terminológica que dificulta seriamente la discusión. ¿Los esquemas de la figura 10.3 son tres tipos de pechinas o bien se trata de una pechina a la izquierda y de dos feas alternativas a la derecha? Como sucede con otros especialistas, los historiadores del arte a menudo no utilizan con el necesario rigor sus términos técnicos. Hablando con precisión, tanto la superficie triangular curvilínea, con su vértice hacia abajo, representada en la figura 10.1, como la superficie representada en la figura 10.3 se denominan en inglés *pendentive* (‘pechina’) y no *spandrel*. Hablando con precisión, un spandrel es lo que queda de una pared cuando se abre un arco redondeado a su través, como se representa en la figura 10.4; es decir, una *enjuta*. Pero incluso esta definición sigue planteando alguna confusión. En la figura 10.4 hemos dibujado sobre un arco redondo, a la izquierda y también a la derecha, lo que es una *enjuta*. ¿Cuentan las «enjutas perforadas» como tales, estrictamente hablando? No lo sé.

Hablando sin tanta precisión, las *enjutas* son espacios que han de ser tratados arquitectónicamente y, en este sentido menos estricto, las tres variaciones de la figura 10.3 cuentan todas como variedades de *enjutas*. Otra variedad de *spandrel* sería la denominada *squinch* (pechina también en castellano), tal como se muestra en la figura 10.5.

Pero a veces, los historiadores del arte emplean el término *enjuta*, cuando en realidad se refieren a *pechina*, la variante dibujada a la izquierda en la

figura 10.3. En este sentido, los *squinchs* no son variantes de *spandrels*, sino sus rivales.

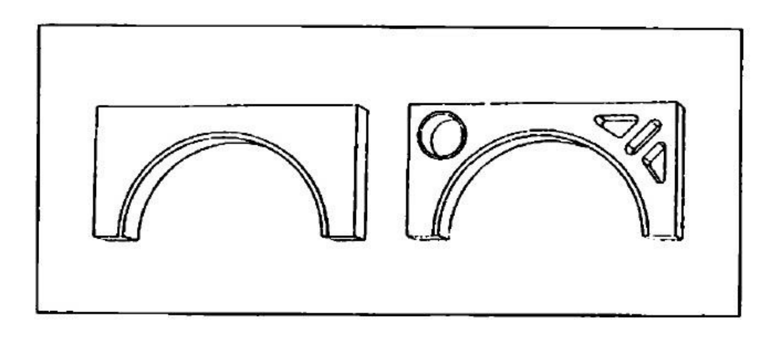


Figura 10.4

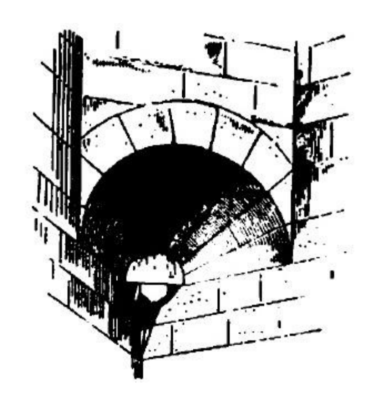


Figura 10.5

Squinch. Un pequeño arco voladizo o nicho a modo de cornisa, colocado entre las paredes que forman las esquinas de una estructura cuadrada para convenirla en una estructura octogonal, de modo que pueda soportar la bóveda de un claustro o una cúpula (Krautheimer 1981).

Entonces, ¿a qué viene todo este lío? Precisamente a que cuando Gould y Lewontin afirman que los *spandrels* son «productos arquitectónicos secundarios y necesarios» lo que dicen es falso, si están utilizando el término *spandrel* en un sentido restringido, como sinónimo de *pechina*, y solamente sería verdadero si entendemos el término *spandrel* en un sentido más amplio que incluya todas las variantes. Pero, en este sentido, los *spandrels* son *problemas* de diseño y no *características* que pueden ser diseñadas (como adaptaciones) o no. Los *spandrels* de Gould y Lewontin, en la versión restringida (sinónimo de *pendentive*, es decir, pechina), son realmente una «necesidad geométrica» en un aspecto, esto es, si tratamos de colocar una

cúpula sobre cuatro arcos, tendremos lo que yo llamo *una oportunidad de diseño obligatoria*: tenemos que colocar allí algo que sostenga la bóveda, sea de una forma o de otra, ya decidiremos cuál. Pero si interpretamos el término *spandrel* como áreas arquitectónicas obligatorias para una u otra adaptación, estos ejemplos tienen dificultades para presentarse como un reto al adaptacionismo.

No obstante, ¿existe alguna otra vía por la cual los *spandrels* de San Marcos, en el sentido restringido del término, es decir, pechinas, sean verdaderamente características no opcionales? Esto es lo que Gould y Lewontin parecen afirmar pero, si es así, están equivocados. No sólo la pechina era una entre otras muchas opciones *imaginables* sino que era justamente una entre otras opciones *disponibles*. Los «squinches» han sido una solución bien conocida en la arquitectura bizantina, aproximadamente desde el siglo VII, para resolver el problema de asentar una bóveda sobre arcos**[100]**.

Con el diseño real de los *spandrels* de San Marcos, es decir, pechinas, se consiguieron básicamente dos cosas. En primer lugar, ésta es, aproximadamente, una superficie denominada de «mínima energía» (la cual se puede obtener si estiramos una película de jabón en un modelo en alambre) y, en consecuencia, cercana a lo que se considera un «área de mínima superficie» (como tal puede ser valorada como la solución óptima si, como se dice, ¡ha de minimizarse el número de costosas piezas de mosaico!). En segundo lugar, esta superficie lisa es *ideal* para colocar sobre ella las imágenes diseñadas con mosaico; y para esto se construyó la basílica de San Marcos: para disponer de un espacio de exposición en el que resplandecieran las imágenes en mosaico. La conclusión es ineludible: los *spandrels* de San Marcos no son pechinas, ni siquiera en el sentido lato de Gould. Son adaptaciones, escogidas de un conjunto de alternativas igualmente posibles, por razones en gran parte estéticas. Han sido *diseñadas* para tener la forma que tienen precisamente con el fin de disponer de superficies apropiadas para desplegar toda la iconografía cristiana.

Después de todo, San Marcos no es un granero; es una iglesia (pero no una catedral). La función primaria de sus bóvedas y de sus cúpulas nunca fue resguardar la iglesia de la lluvia —ya que cuando las bóvedas fueron construidas en el siglo XI existían otras formas menos costosas de hacerlo—, sino disponer de áreas para la exposición de los símbolos del credo. Sobre el mismo asentamiento de San Marcos se había quemado una iglesia que fue reconstruida en el año 976, aunque posteriormente el estilo bizantino de la

decoración con mosaicos había llamado la atención de los poderosos venecianos, que desearon crear un ejemplo en su ciudad. Otto Demus, [1984], la gran autoridad en los mosaicos de San Marcos, demostró, en cuatro magníficos volúmenes, que los mosaicos son la *raison d'être* de San Marcos y, en consecuencia, de muchos de sus detalles arquitectónicos. En otras palabras, no habría tales pechinas en Venecia si no se hubiese presentado y encontrado la solución al «problema del medio ambiente», en concreto a la disposición de las imágenes en mosaico bizantino de la iconografía cristiana. Si observamos de cerca las pechinas (esto se detecta en la figura 10.1 pero muchísimo más si miramos las pechinas reales, como tuve oportunidad de hacer en una reciente visita a Venecia) observaremos la atención que se dedicó para conseguir suavizar la transición entre la pechina propiamente y los arcos con los que se conecta de la mejor manera posible para disponer de una superficie continua en la que aplicar los mosaicos.

El otro ejemplo arquitectónico de Gould y Lewontin, la capilla del King's College, también está mal elegido, como se ha demostrado, ya que no sabemos si las decoraciones alternantes, rosa de Tudor o rastrillo, son la *raison d'être* de la bóveda en abanico o viceversa. Sabemos que la bóveda en abanico no era parte del diseño original de esta capilla, pero que en una revisión posterior se introdujo un cambio, años después de que la construcción hubiera comenzado y por razones desconocidas (Fitchen 1961:248). Las muy pesadas piedras claves (y muy esculpidas), colocadas en las intersecciones de los nervios de las primitivas bóvedas góticas, habían sido para los constructores algo así como un «movimiento forzado», como ya he señalado en el capítulo 8, dado que necesitaban el peso extra de las piedras claves para contrarrestar la tendencia a elevarse de los arcos puntiagudos, especialmente durante la fase de construcción, cuando la deformación de las estructuras parcialmente construidas constituía un gran problema que necesitaba solución. Pero en las últimas bóvedas en abanico como la del King's College, el propósito de los espacios donde se producía la intersección de los nervios era, probablemente, disponer de superficies para la ornamentación. ¿Han de estar en esta posición las áreas de intersección de los nervios? No. Desde el punto de vista de la ingeniería podría haber existido en esa situación orificios redondeados, «linternas» que dejaran penetrar la luz del día desde arriba, de no estar en el techo. ¡Pudiera ser que la bóveda en abanico *fuera* escogida por los constructores para que el techo de la capilla incorporara los símbolos de los Tudor!

Así que, después de todo, los tabulados *spandrels* de San Marcos no son *spandrels*, sino adaptaciones[101] Lo cual no deja de ser curioso, aunque no teóricamente importante, debido a que, como el propio Gould nos ha recordado, uno de los mensajes fundamentales de Darwin es que los artefactos pueden reciclarse para cumplir nuevas funciones: *exapted* según el neologismo inglés acuñado por Gould y Vrba [1981]. El pulgar del panda no es realmente un pulgar, aunque es muy bueno haciendo lo que hace. ¿No es acaso el concepto de Gould y Lewontin de la pechina una herramienta valiosa en el pensamiento evolucionista aunque su nacimiento fuera, reciclando otra famosa frase, un accidente congelado de la historia? En definitiva, ¿cuál es la función del término *spandrel* en el pensamiento evolucionista? Hasta donde yo sé, Gould nunca le ha otorgado al término (en su aplicación a la biología) una definición oficial, y dado que los ejemplos que él ha utilizado para mostrar el pretendido significado son, cuando menos, equívocos, estamos obligados a utilizar nuestros propios recursos: debemos encontrar la mejor, la más caritativa interpretación de sus textos. Cuando volvemos a la tarea, emerge claramente del contexto una cuestión: sea lo que sea una pechina se supone que no es una adaptación.

¿Cuál sería un *buen* ejemplo arquitectónico de pechina en el sentido de Gould? Si las adaptaciones son ejemplos de diseño —bueno y hábil—, quizás entonces una pechina es algo no inteligente, una característica que no exhibe, en modo alguno, una habilidad en el diseño. La existencia del vano de una puerta —simplemente una abertura— en un edificio podría ser un ejemplo, ya que no nos impresiona particularmente la sabiduría del constructor que incluye en su casa tal característica. Pero existe, después de todo, una muy buena razón por la que los edificios tienen vanos de puertas. Si las pechinas son soluciones totalmente obvias para problemas de diseño que, por lo tanto, tienden a formar parte de una tradición inconsciente de la construcción de edificios, entonces los *spandrels* abundan. Sin embargo, en este caso las pechinas no serían alternativas a la adaptación sino ejemplos *por excelencia* de adaptación, ya se trate de «movimientos forzados» o, en cualquier caso, movimientos que sería una necedad no tener en cuenta. Un ejemplo más apropiado podría ser lo que los ingenieros llaman a veces un «no se preocupe»: algo que se conseguirá de un modo o de otro, pero sin que lo que se consiga sea mejor de un modo o de otro. Si colocamos una puerta en un vano, la puerta necesitará bisagras, pero ¿dónde deben ir colocadas, a la derecha o a la izquierda? Quizás a nadie le importe, así que lanzamos una moneda al aire y las bisagras quedan instaladas a la izquierda. Si otros

constructores copian automáticamente los resultados, sin pensarlo, están estableciendo una tradición local (reforzada por los fabricantes de picaportes, que hacen picaportes para las puertas con bisagras a la izquierda), el resultado puede ser una pechina simulando una adaptación. «¿Porqué tienen todas las puertas de este pueblo las bisagras colocadas a la izquierda?» sería una clásica pregunta adaptacionista, para la cual la respuesta sería: «No hay razón alguna. Se trata sencillamente de un accidente histórico». ¿Es éste un buen ejemplo arquitectónico de una pechina? Quizá, pero, como sucede en el ejemplo de las hojas de otoño expuesto en el capítulo precedente, nunca es un error utilizar los adaptacionistas «porqués», aunque la respuesta verdadera sea que no hay razón para ello. ¿Hay muchas características en la biosfera que existan sin razón alguna? Todo depende de lo que entendamos por característica. De un modo trivial, hay muchas propiedades (por ejemplo, la propiedad del elefante de tener más extremidades que ojos, la flotabilidad de las margaritas) que no son, en sí mismas, adaptaciones, pero que ningún adaptacionista negaría. Presumiblemente existe una doctrina más interesante que Gould y Lewontin nos apremian a que abandonemos.

¿Cuál es la doctrina de la «adaptación invasiva», aquella que Gould supuso que una admisión de pechinas ampliamente extendidas debería destruir? Consideremos la más extrema forma de adaptacionismo panglossiano imaginable, es decir, aquella que sostuviera que toda *cosa diseñada* está *óptimamente* diseñada. Una mirada de reojo a la ingeniería humana mostrará que incluso esta visión no sólo permite, sino que requiere la existencia de una multitud de cosas no diseñadas. Imaginemos, si es posible, alguna obra maestra de la ingeniería humana: una fábrica de artefactos perfectamente diseñada, eficiente en el consumo de energía, máximamente productiva, mínimamente costosa de operar, máximamente humana con los trabajadores, sencillamente a toda prueba en cualquier dimensión. El sistema de recogida de papel usado, por ejemplo, reciclaría hacia el tipo de papel máximamente conveniente y agradable para el personal, al mínimo coste energético, y así sucesivamente. Un triunfo panglossiano, según parece. Pero, un momento, ¿para qué es el papel usado? No es para algo. Es un producto secundario de otros procesos y el sistema de recogida de papel usado es para tratarlo. ¡No podemos dar una explicación adaptacionista de por qué el sistema de recogida y reciclado es óptimo sin presuponer que el papel usado, en sí mismo, es justamente… usado! Naturalmente, podemos ir y preguntar si las labores de los oficinistas pueden realizarse «sin papeles» con una mejor utilización de los ordenadores, pero si esto sucede, sea por una razón o por

otra, quedaría aún el papel usado para ser tratado y otras cosas usadas y productos secundarios en cualquier caso, de modo que habría siempre multitud de características no diseñadas en un sistema que está máximamente bien diseñado. Ningún adaptacionista podría ser un adaptacionista tan «invasivo» como para negar esto. Por lo que sé, la tesis de que toda propiedad, de que toda característica de lo que se encuentra en el mundo viviente es una adaptación, no es una tesis que ninguno haya tomado en serio, o haya sido sugerida por alguien que fuera tomado en serio. Si no me equivoco, hay algunos estúpidos, pero Gould nunca ha demostrado ser uno de ellos.

A veces, sin embargo, parece que Gould cree que éste es el modo de atacar. Gould caracteriza el adaptacionismo como «puro adaptacionismo» y «panadaptacionismo»; aparentemente el punto de vista que sostiene que toda característica de cualquier organismo ha de explicarse como seleccionada para la adaptación. En su reciente libro *The Ant and the Peacock* (1991), Helena Cronin, dedicada a la filosofía de la biología, es especialmente aguda en su diagnóstico. Ella descubre a Gould en el acto de caer precisamente en la siguiente interpretación incorrecta:

> Stephen Gould habla acerca de «lo que puede ser la cuestión más fundamental en la teoría evolucionista» y en ese momento, de modo significativo, no plantea una cuestión sino dos: «¿Hasta qué punto es *exclusiva* la selección natural como un agente del cambio evolutivo? ¿Deben considerarse como adaptaciones *todas* las características de los organismos?» (Gould 1980a:49; la cursiva es mía). Pero la selección natural puede ser el único verdadero generador de adaptaciones sin haber generado todas las características; es posible sostener que todas las características adaptan vas son el resultado de la selección natural sin sostener por ello que todas las características sean, realmente, adaptativas (Cronin 1991:86).

La selección natural puede ser, en consecuencia, el «agente exclusivo» de los cambios evolutivos aun cuando muchas características de los organismos no sean adaptativas. Los adaptacionistas están —y deben estar— *siempre* alertas, dispuestos a dar explicaciones adaptativas de cualquier característica que llame su atención, pero esta estrategia corre el riesgo de convertir a alguno en la caricatura de lo que Gould llama «panadaptacionismo».

Quizá quedaría más claro a qué se opone Gould, si prestamos atención a lo que recomienda. ¿Qué alternativas al adaptacionismo, como componentes de su recomendado pluralismo, sugieren Gould y Lewontin? La principal es la idea del *Bauplan*, un término arquitectónico alemán adoptado por algunos biólogos europeos. *Bauplan* se traduce habitualmente como plano básico o diseño de planta de una estructura, vista desde arriba. Es curioso que en una campaña *anti*-adaptacionista ocupe un lugar destacado un término

arquitectónico, aunque parece una estupidez cuando comprobamos cómo lo introdujeron los teóricos del original *Bauplan*. La adaptación, dicen Gould y Lewontin, puede explicar las modificaciones *superficiales* del diseño del organismo para adaptarse al medio ambiente, pero no las características fundamentales de los seres vivos: «Los pasos importantes de la evolución, la construcción del propio *Bauplan* y la transición entre los diversos planos básicos, debe implicar a algún otro mecanismo desconocido, quizás "interno"» Gould y Lewontin 1979:159). ¿Están diciendo Gould y Lewontin que el *Bauplan* no está diseñado por la evolución sino que se trata de algo dado? Esto suena como algo inverosímil, ¿no les parece? ¿Estaban Gould y Lewontin importando de Europa esta idea radical? No, por el momento. Rápidamente dan por probado que los biólogos ingleses han actuado correctamente «al rechazar esta idea, tan cercana a una llamada al misticismo» (Gould y Lewontin 1979:159).

Pero una vez que la versión mística del *Bauplan* ha sido soslayada, ¿qué queda? Nuestra vieja amiga: la afirmación de que la buena ingeniería revertida roma en sus manos el proceso de construcción. Como puntualizan Gould y Lewontin, sus puntos de vista sobre este asunto «no niegan que el cambio, cuando ocurre, pueda ser mediado por la selección natural, aunque mantienen que hay limitaciones que restringen las posibles vías y modos de cambio con tanta fuerza que dichas limitaciones se convierten en uno de los aspectos más interesantes de la evolución» (1979:160). Sean o no las limitaciones uno de los aspectos más interesantes, lo cierto es que, como hemos visto, son importantes. Quizá los adaptacionistas, al igual que los historiadores del arte, necesiten dirigir repetidamente su atención a este punto importante. Cuando Dawkins, un archiadaptacionista donde los haya, dice: «Hay algunas formas que, desde ciertos tipos de desarrollo embriológico, parecen incapaces de crecimiento» (Dawkins 1989b:216), está expresando una versión de este punto acerca de las limitaciones del *Bauplan*, lo cual, según dice, fue algo así como una revelación para él. Lo consiguió esforzadamente a través de sus propias simulaciones de la evolución en el ordenador, no a través del artículo de Gould y Lewontin, aunque les podemos permitir que tercien en la discusión diciendo: «¡Ya lo decíamos nosotros!».

Gould y Lewontin también discuten otras alternativas a la adaptación, y éstas, por lo demás, son temas que ya hemos encontrado en el darwinismo ortodoxo: fijación aleatorizada de genes (el papel del accidente histórico y su amplificación), limitaciones del desarrollo debidas al modo según el cual se expresan los genes y los problemas de moverse en un «paisaje adaptativo»

con «múltiples picos». Todos estos son fenómenos reales; como es habitual, el debate entre los evolucionistas no es sobre si existen estos fenómenos, sino hasta qué punto son importantes. Las teorías que los incorporan han desempeñado realmente un papel significativo dentro de la creciente sofisticación de la síntesis neodarwiniana, pero son reformas o complicaciones, no revoluciones.

De este modo, algunos evolucionistas han aceptado el pluralismo de Gould y Lewontin con un espíritu irenista, no como una llamada al abandono del adaptacionismo, sino a su mejoramiento. Como Maynard Smith ha escrito: «El efecto de la publicación de Gould y Lewontin ha sido considerable y, en conjunto, bienvenido. Dudo de que mucha gente se haya cansado leyendo el relato de las historias de las adaptaciones. Ciertamente, a mí no me sucedió» (1991:6). El artículo de Gould y Lewontin ha tenido un efecto que ha sido bienvenido, pero uno de sus productos secundarios no ha sido tan bien recibido. La inflamada retórica de Gould y Lewontin sugiriendo que estos temas algo olvidados constituyen una alternativa mayor al adaptacionismo, abrieron las compuertas a una gran oleada de pensamiento voluntarista generada por los que temen a Darwin, aquellos que preferirían que no hubiese una explicación adaptacionista para ningún preciado fenómeno. ¿Cuál sería la imaginada y nebulosa alternativa? La «necesidad interna», que los propios Gould y Lewontin descartan como una apelación al misticismo o una absoluta coincidencia cósmica, un no-comienzo igualmente místico. Ni Gould ni Lewontin apoyan explícitamente una alternativa verdaderamente nueva a la adaptación, aunque esta actitud haya pasado inadvertida para aquellos que desean ser deslumbrados por la autoridad de estos eminentes incrédulos del pensamiento de Darwin.

Además, Gould, a pesar de su llamada al pluralismo, en el artículo firmado en colaboración con Lewontin [1993a], ha persistido en describirlo como terreno baldío para el adaptacionismo y ha mantenido una interpretación no darwiniana de su concepto central, las pechinas. Puede haber sucedido que me haya pasado por alto una interpretación obvia de las pechinas: quizás éstas sean, precisamente, un fenómeno QWERTY. Los fenómenos QWERTY, como recordaréis, son limitaciones, pero limitaciones con una historia adaptativa y, por tanto, con una explicación adaptacionista**[102]**. El propio Gould considera brevemente esta alternativa: «Si los canales (los que limitan las opciones actuales) están ahí a partir de pasadas adaptaciones, entonces la selección sigue siendo preeminente, y todas las estructuras mayores son expresión de una selección inmediata o han sido

canalizadas por la herencia filogénica de una selección previa» (1982a:383). Una bella exposición, aunque enseguida la rechazó, denominándola darwiniana, lo cual es cierto, y recomendó una «versión no darwiniana» alternativa, que describe como «no ampliamente apreciada, pero potencialmente fundamental». Las pechinas, sugiere Gould, son limitaciones congeladas creadas por adaptaciones más tempranas: son ex-(ad)aptaciones. ¿Qué contraste estaba tratando de describir Gould?

Creo que él se dio cuenta de la diferencia entre el aprovechamiento de algo previamente diseñado y el aprovechamiento de algo que en su origen fue insuficientemente diseñado, y decidió sostener que era una diferencia importante. Quizá. He aquí una evidencia textual indirecta de esta lectura. Un reciente artículo publicado en el Boston Globe cita al lingüista Samuel Jay Keyser del MIT:

«El lenguaje puede compararse con una pechina de la mente», dice Keyser y espera pacientemente mientras su interlocutor busca en el diccionario la palabra pechina. «El primer constructor que apoyó cúpulas sobre arcos creó las pechinas por *accidente* y los primeros constructores no prestaron atención a las pechinas y decoraban únicamente los arcos», dice Keyser. «Pero después de un par de siglos los constructores comenzaron a prestar atención y a decorar las pechinas. Del mismo modo», dice Keyser, «el lenguaje —es decir, la habilidad para transmitir información mediante el habla— pudo haber sido una pechina del pensamiento y la comunicación creada accidentalmente por el desarrollo de algún “arco” cultural…». «El lenguaje es, muy probablemente, un artefacto accidental derivado de alguna peculiaridad evolutiva de la mente» (Robb 1991). (La cursiva es mía).

Quizás a Keyser lo han citado erróneamente; siempre soy muy cauto cuando se trata de aceptar el relato de un periodista acerca de las palabras de alguien, ya que lo he sufrido en mis carnes, pero si la cita es exacta, entonces para Keyser las pechinas son originalmente accidentes, no necesidades, «algo que da lo mismo» o fenómenos QWERTY. En cierta ocasión, cuando estaba trabajando en una fundición de moldes de bronces en Roma, se produjo una explosión en uno de los moldes que estábamos rellenando: el bronce líquido se esparció por el suelo. Parte de la masa esparcida del bronce se endureció adquiriendo la forma de un fantástico encaje que enseguida recogí y lo convertí en una escultura. ¿Estaba yo «*ex*-adaptando» una pechina? (Marcel Duchamp, el dadaísta, por el contrario, no estaba *ex*-adaptando una pechina cuando se apropió de un orinal como su *objet trouvé* y lo llamó escultura, dado que el orinal en su primera vida cumplía una función).

El propio Gould [1993:31] ha citado con aprobación esta historia del *Boston Globe* sin darse cuenta de que la historia del arte relatada por Keyser estaba equivocada, y sin expresar ningún desacuerdo con la definición de pechina de Keyser, considerada como un accidente. Así, quizá Keyser esté en

lo cierto sobre el significado del término: las pechinas son justamente accidentes disponibles para «ex-(ad)aptación». Gould introdujo el término inglés *exaptation* en un artículo escrito en colaboración con Elizabeth Vrba en 1982 y titulado: «Exaptation: A Missing Term in the Science of the Form» (‘Ex-(ad)aptación: Un término olvidado en la ciencia de las formas’). Su intento fue comparar ex-(ad)aptación con adaptación. Sin embargo, su principal candidato falso era un término increíblemente desfavorecido que había conseguido alguna circulación en los libros de textos sobre la evolución, la *preadaptación*.

Preadaptación parece implicar que la protoala, mientras comenzaba con sus primeros escarceos funcionales, en sus estadios incipientes, sabía ya hacia donde se dirigía: estaba predestinada para una posterior conversión a la función de volar. Los libros de texto introducen habitualmente la palabra preadaptación y entonces rápidamente rechazan cualquier aroma a ordenación previa. (Pero un vocablo ha sido obviamente mal elegido si no puede ser utilizado sin negar su significado literal) (Gould 1991b:144n).

«Preadaptación» era un término terrible, exactamente por las razones que expone Gould, pero nótese que él no está afirmando que los objetivos de sus críticas se dirijan al gran error de otorgar previsión a la selección natural; él admite que «rápidamente rechazan» esta herejía, en el mismo momento de introducir el término. Están cometiendo el error menor de elegir perversamente su uso para fomentar esta confusión. El paso de «preadaptación» a «ex-(ad)aptación» puede ser muy bien visto como una sabia reforma del uso, más apropiada para apropiarse del punto de vista ortodoxo de los adaptaciónistas. Gould, sin embargo, resistió esta interpretación reformista. Deseaba los términos «ex-(ad)aptación» y «pechinas» para presentar una alternativa «potencialmente fundamental» y «no darwiniana».

Elisabeth Vrba y yo hemos propuesto que el restrictivo y confuso término «preadaptación» sea eliminado en favor del término más completo de «ex-(ad)aptación» —para cualquier órgano no evolucionado bajo la selección natural para su uso actual—, sea porque realiza una función diferente de la de sus ancestros (preadaptación clásica) o porque representa una parte no funcional disponible para una posterior cooptación (Gould 1991b:144n).

Pero, de acuerdo con el darwinismo ortodoxo, toda adaptación es una forma u otra de ex-(ad)aptación; esto es trivial, dado que la función no es eterna; si retrocedemos lo suficiente encontraremos que toda adaptación se ha desarrollado a partir de estructuras del predecesor, cada una de las cuales tenían algún otro uso, o ninguno. Los únicos fenómenos que la revolución de la ex-(ad)aptación de Gould descarta son los que los adaptacionistas

ortodoxos desaprueban «rápidamente», en cualquier caso: las preadaptaciones planeadas para algo.

La revolución de la pechina (contra el panadaptacionismo) y la revolución de la ex-(adaptación (contra el preadaptacionismo) se evaporan ante una inspección más de cerca, dado que ambos, panadaptacionismo y preadaptacionismo, han sido habitualmente rehuidos por los darwinistas, incluyendo al propio Darwin. Estas no revoluciones no sólo no suponen un reto para ningún principio darwiniano ortodoxo: las acuñaciones que introducen confundirán tanto como las acuñaciones con las que pretendían reemplazarlas.

Es duro ser un revolucionario si el *establishment* mantiene con aquel una postura de cooptación. Gould se ha quejado, a menudo, de que su objetivo, el neodarwinismo, reconoce las excepciones que él desea convertir en objeciones y «esto impone una gran frustración a todo aquel que caracteriza la síntesis darwiniana moderna con el propósito de criticarla» (Gould 1980:130).

La síntesis moderna ha sido a veces construida con tanta amplitud, por lo general, por defensores que desean verla completamente apropiada para encontrarse y abarcar las críticas actuales, que pierde todo significado al incluirlo todo (…) Stebbins y Avala (dos eminentes defensores) han tratado de ganar un argumentó por redefinición. La esencia de la síntesis moderna debe ser su núcleo darwiniano (Gould 1982a:382).

Es sorprendente ver a un darwiniano concediéndole a algo una esencia, pero podemos apropiarnos del punto de vista de Gould, aunque no de su lenguaje: hay algo acerca de la moderna síntesis que Gould desea derribar, y antes de poder derribar algo, ha de ser inmovilizado. Gould [1983a] ha afirmado, que se imaginaba la moderna síntesis haciéndole el trabajo, al «endurecerse» con una quebradiza ortodoxia que sería más fácil de atacar. ¡Si sólo fuera esto! De hecho tan pronto como Gould se ha lanzado a la batalla, la síntesis darwiniana moderna ha demostrado su flexibilidad, absorbiendo sus golpes, para su frustración. Yo creo que Gould está en lo cierto cuando dice que la síntesis moderna tiene un «núcleo darwiniano», y también creo que está en lo cierto cuando asegura que éste es su objetivo; pero Gould todavía no ha puesto el dedo en ese núcleo.

Si el proceso abierto contra la «adaptación invasiva» se ha desvanecido, entonces ¿qué hay del proceso contra el gradualismo, el otro elemento principal de la moderna síntesis que Gould imagina en pleno desmoronamiento? La pretendida revolución de Gould contra el gradualismo fue en realidad la primera; la inició con una salva en 1972 en la cual introdujo

otra acuñación familiar al vocabulario de los evolucionistas y de los espectadores: el *equilibrio puntuado*.

3. *Equilibrio puntuado: un monstruo prometedor*

> El equilibrio puntuado ha conseguido finalmente una mayoría de edad nada ambigua, es decir, nuestra teoría tiene ahora 21 años. Nosotros también, con orgullo paterno (y sin embargo, admitiendo los sesgos potenciales) creemos que la controversia primaria ha cedido el paso a una comprensión general y que el equilibrio puntuado ha sido aceptado por la mayoría de nuestros colegas (un tipo de mayoría más convencional) como una valiosa adición a la teoría de la evolución.
>
> Stephen Jay Gould y Niles Eldredge
> *«Punctuated Equilibrium comes of Age»*

> Lo que es necesario decir ahora, alto y claro, es la verdad: que la teoría del equilibrio puntuado se asienta firmemente dentro de la síntesis neodarwinista. Siempre fue así. Tomará tiempo reparar el daño provocado por la arrolladora retórica, pero será reparado.
>
> Richard Dawkins, *El relojero ciego*

Niles Eldredge y Gould colaboraron en el artículo que introdujo el término: «El equilibrio puntuado. Una alternativa al gradualismo filético» (1972). Mientras que, según los autores del artículo, los darwinistas ortodoxos tendían a considerar todo cambio evolucionista como gradual, ellos argumentaban que, por el contrario, la evolución procedía mediante saltos; largos períodos sin cambios, denominados *estasis* —equilibrio—, interrumpidos por períodos breves, súbitos y dramáticos, de rápidos cambios denominados puntuaciones. La idea básica se ilustra habitualmente comparando un par de árboles de la vida (figura 10.6).

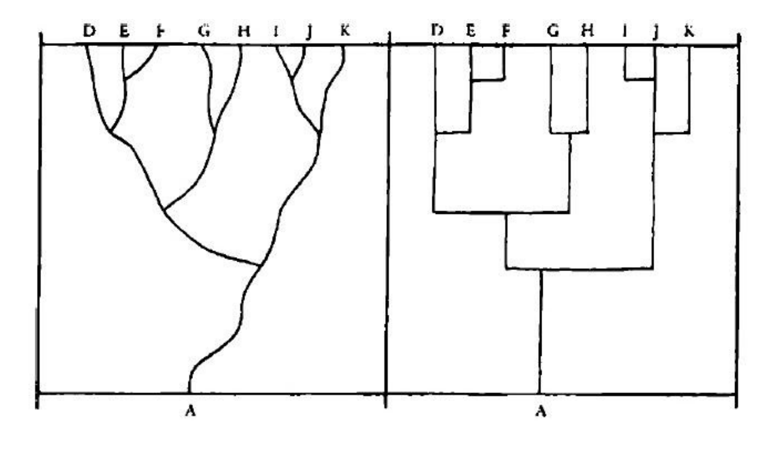


Figura 10.6

Podemos suponer que la dimensión horizontal en el diagrama registra algún aspecto de variación fenotípica o diseño corporal; necesitaríamos, naturalmente, un espacio multidimensional para representar al equilibrio puntuado en su totalidad. El punto de vista ortodoxo, en el cuadrado de la izquierda de la figura, se ha dibujado mostrando que todos los movimientos a través del espacio de diseño (esto es, a la izquierda o a la derecha del diagrama) siguen un ritmo más o menos sostenido. El equilibrio puntuado, por el contrario, muestra largos períodos de diseño no modificado (los segmentos verticales de las líneas) interrumpidos por desviaciones laterales «instantáneas» en el espacio de diseño (los segmentos horizontales). Para conocer la tesis central de esta teoría, sígase la historia evolucionaría de las especies hasta *K* en cada dibujo. El dibujo ortodoxo, a la izquierda, muestra una desviación más o menos sostenida hacia la derecha en el diagrama desde el Adán de las especies, A. La propuesta alternativa está de acuerdo en que *K* es un descendiente de A y que esto se ha conseguido con la misma desviación hacia la derecha en el espacio de diseño, en la misma cantidad de tiempo, pero con paradas y comienzos, no como una escalada mantenida. (Estos diagramas pueden ser engañosos; la diferencia entre una rampa y una escalera es el punto de contraste, pero los pasos gigantes son los movimientos laterales, no los *bits* verticales, que son los aburridos períodos de «movimiento» sólo a través del tiempo, sin movimiento a través del espacio de diseño).

Frente a una hipótesis con propósito de radicalidad suele producirse un familiar trío de reacciones en los científicos: a) «¡Debe de estar loco!», b) «¿Qué tiene de nuevo? ¡Todo el mundo conoce es/o!», y, más tarde —si la hipótesis resiste—, c) «¡Vaya! ¡Aquí hay gato encerrado!». A veces estas

fases tardan años en desarrollarse, una tras otra, aunque he visto emerger a las tres en casi total sincronía, durante el transcurso de media hora de acalorada discusión, tras una conferencia dedicada a este artículo de Gould. En el caso de la hipótesis del equilibrio puntuado, las fases se reiteran más, pronunciadas, en gran parte debido a que Gould ha cambiado varias veces su pensamiento acerca de lo que él y Eldredge estaban afirmando. En su primera aparición, la tesis del equilibrio puntuado no se presentó como un reto revolucionario, sino como una corrección conservadora de una ilusión a la que habían sucumbido los darwinistas ortodoxos: los paleontólogos estaban sencillamente equivocados cuando pensaban que la selección natural darwiniana debía dejar una memoria física de fósiles que mostrara multitud de formas intermedias[103]. En el primer artículo de Gould y Eldredge no se menciona ninguna teoría radical de especiación o mutación. Pero más tarde, alrededor de 1980, Gould decidió que el equilibrio puntuado, después de todo, era una idea revolucionaria, no una explicación de la ausencia de gradualismo en la memoria física de los fósiles, sino una refutación del propio gradualismo darwiniano. Esta declaración se anunció como revolucionaria y, en esta ocasión, en verdad lo era. Demasiado revolucionaria, fue abucheada con la ferocidad que el *establishment* reserva a los herejes como Elaine Morgan. Gould dio marcha atrás enérgicamente, negando repetidas veces que él hubiera querido decir algo tan extravagante. En ese caso, respondió el *establishment*, no hay nada nuevo en lo que Gould dice. Pero ¡un momento! ¿Puede existir alguna otra lectura de la hipótesis, según la cual la teoría sea verdadera y nueva? Es posible. La fase tres está aún en marcha y el jurado se había retirado a deliberar varias alternativas, aunque no todas no revolucionarias. Tenemos que volver a seguir las huellas de las fases para ver por qué se ha producido todo este alboroto.

Como los mismos Gould y Eldredge han señalado, existe un obvio problema de escala en los dibujos de la figura 10.6. ¿Qué pasaría si aumentáramos el diagrama ortodoxo mediante un zoom y encontrásemos que lo que vemos aparece como se representa en la figura 10.7?

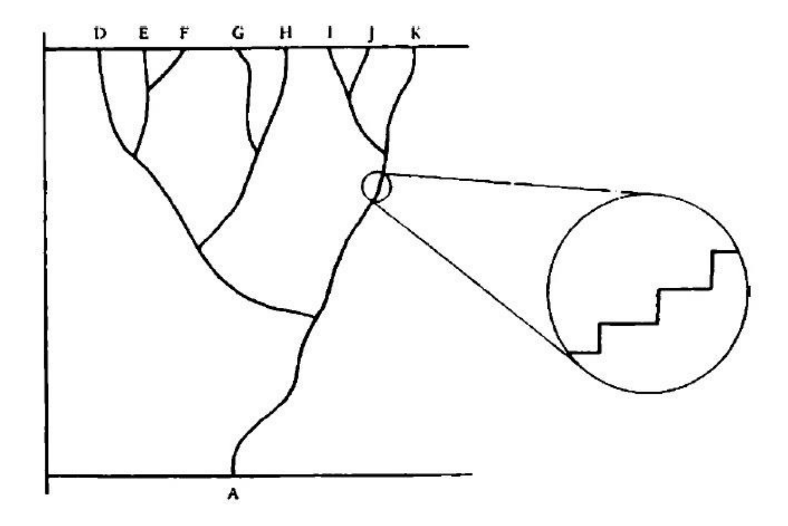


Figura 10.7

A algún nivel de ampliación cualquier rampa evolutiva debe aparecer como una escalera. ¿Es la figura 10.7 un diagrama del equilibrio puntuado? Si lo es, entonces el darwinismo ortodoxo era ya una teoría de equilibrio puntuado. Incluso el más radical gradualista puede permitir que la evolución se tome un respiro de vez en cuando, dejando que las líneas verticales se extiendan indefinidamente a través del tiempo hasta que surja alguna nueva presión de selección, de un modo o de otro. Durante este período de estasis, la presión de selección debe ser conservadora, manteniendo el diseño aproximadamente constante mediante la rápida eliminación de cualquier alternativa experimental que aparezca. Como dicen los mecánicos viejos, «No repares lo que no esté roto». Dondequiera que surge una nueva presión de selección veremos una «súbita» respuesta de incremento de la evolución, una puntuación que interrumpe el equilibrio. ¿Había sido realmente presentado por Eldredge y Gould un punto de desacuerdo ciertamente revolucionario o simplemente se trataba de una observación interesante acerca de la variabilidad en el *tempo* de los procesos evolutivos y sus previsibles efectos en el registro de los fósiles?

Es característico de los puntualistas dibujar en sus revolucionarios diagramas los segmentos correspondientes a la puntuación como absolutamente horizontales (para que quede muy claro que están presentando una verdadera alternativa al punto de vista de la ortodoxia, con sus rampantes rampas). Esto hace que las revisiones de los diseños aparezcan como si tuvieran lugar en un abrir y cerrar de ojos, en un dos por tres. Pero precisamente se trata de un artefacto engañoso de la enorme escala vertical

adoptada, que representa millones de años en una pulgada. Sin embargo el movimiento lateral de la llamada «puntuación» no es realmente instantáneo, sino tan sólo «geológicamente instantáneo».

Una población aislada puede tardar mil años en formar una especie y por lo tanto, su transformación puede parecer glacialmente lenta, si se mide con la irrelevante escala de nuestras vidas personales. Pero mil años, registrados apropiadamente en el tiempo geológico, es solamente un momento indisoluble, generalmente preservado en un único plano estratificado (en una roca que contiene fósiles), en el tiempo vital de especies que, a menudo, viven en estasis durante varios millones de años (Gould 1992a:12-14).

Supongamos que, utilizando un zoom, hemos ampliado la imagen de uno de estos instantes de miles de años, cambiando la escala vertical de la dimensión del tiempo en unos cuantos órdenes de magnitud, para observar lo que realmente pueda estar sucediendo (figura 10.8). El paso horizontal realizado entre el tiempo t y el tiempo t’ habrá de ser algo alargado, y debemos convertirlo relativamente en pasos grandes o pasos pequeños o pasos diminutos, o en una combinación de éstos.

¿Eran revolucionarias cualquiera de estas posibilidades? ¿Qué tesis mantenían exactamente Eldredge y Gould? En esta cuestión sus respectivos puntos de vista divergían algo, al menos durante algún tiempo. La tesis era revolucionaria, afirmaba Gould, debido a que sostenía que las puntuaciones no eran precisamente la evolución habitual, no se trataba de cambios *graduales*. Recordemos el viejo chiste del borracho que se cae por la escalera mecánica y dice, mientras sube: «¡Atentos al primer escalón! ¡Es el mejor en su clase!». Durante algún tiempo, Gould propuso que el primer paso en el establecimiento de cualquier nueva especie era el mejor en su clase; un *salto* no darwiniano.

La especiación no es siempre una extensión de una gradual y adaptativa sustitución alélica para conseguir un efecto mayor, aunque puede representar, como Goldschmidt argumenta, un estilo diferente de cambio genético, tal como una rápida reorganización del genoma, quizá no adaptativa (Gould 1980b:119).

La propia especiación, desde este punto de vista, no es el efecto de adaptaciones acumuladas gradualmente que diferencian poblaciones, sino, más bien, una causa con su propia explicación no darwiniana.

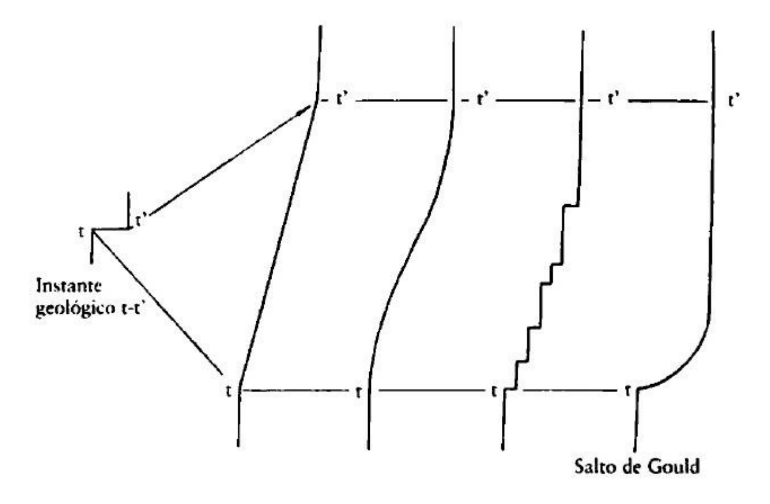


Figura 10.8

Pero en la especiación cromosómica, que da saltos, el aislamiento reproductivo es lo primero y no puede ser considerado en absoluto como una adaptación… Podemos, de hecho, invertir el punto de vista convencional y argumentar que la especiación, formando nuevas entidades de modo estocástico, aporta materia prima para la selección (Gould 1980b:12.4).

Esta sugerencia, que yo llamo el salto de Gould, está representada a la derecha en la figura 10.8. Sólo parte del proceso de la puntuación, un gradual proceso de limpieza al final, puede considerarse «darwiniano», afirma Gould:

Si surgen nuevos planos corporales básicos (*Baupläne*) en una cascada adaptativa, que ha tenido su origen a partir de un salto en una característica clave, entonces puede decirse que parte del proceso es secuencial y adaptativo, y por lo tanto darwiniano; pero el paso inicial no lo es, dado que la selección no desempeña un papel creativo en la construcción de la característica clave (Gould 1982a:383).

Es este «papel creativo» de algo que no era la selección lo que captó la escéptica atención de los colegas de Gould. Para esclarecer qué causó este furor, es necesario hacer notar que nuestro diagrama en la figura 10.8 es realmente incapaz de distinguir varias hipótesis crucialmente diferentes. El problema con el diagrama es que necesita más dimensiones, de modo que podamos comparar los pasos en el *espacio del genotipo* (los pasos tipográficos en la Biblioteca de Mendel) con los pasos en el *espacio del fenotipo* (las innovaciones de diseño en el espacio de diseño) y, entonces, evaluar estas diferencias en un *paisaje adaptativo*. Como hemos visto, las relaciones entre la receta y el resultado son complejas, y pueden ilustrarse muchas posibilidades. Vimos en el capítulo 5 que un pequeño cambio tipográfico en el genoma podría, en principio, provocar un gran efecto en el

fenotipo expresado. Comprobamos también, en el capítulo 8, que algunos cambios tipográficos en el genoma pueden no provocar efecto alguno en el fenotipo: hay más de cien formas diferentes de «deletrear» lisozima, por ejemplo, y en consecuencia, más de cien formas equivalentes de deletrear el orden para la lisozima en los codones de ADN. Sabemos, entonces, que en un extremo puede haber organismos tan similares en el diseño como para ser indistinguibles, que no obstante tienen grandes diferencias en su ADN; por ejemplo, uno de nosotros y cualquier otra persona con la que a menudo somos confundidos (el doble, *Doppelgänger* en alemán; ¡ningún libro de filosofía sería completo si no mencionara a los dobles!). En el otro extremo, puede tratarse de organismos que son extrañamente diferentes en apariencia pero que son casi idénticos genéticamente. Una simple mutación justamente en el lugar equivocado puede producir un monstruo; el término médico para tales progenies deformadas es *terata*, vocablo griego (y latino) para designar los «monstruos». Y también pueden haber organismos que son casi idénticos en apariencia y estructura, y casi idénticos en el ADN, pero tremendamente diferentes en su *capacidad de adaptación al medio*; por ejemplo, los casos de hermanos gemelos, uno de los cuales tiene un gen que le aporta inmunidad o susceptibilidad con respecto a alguna enfermedad.

Un gran saltó en cualquiera de estos tres espacios puede también denominarse una *macromutación* (con el significado de una gran mutación, no precisamente una mutación en lo que he llamado un *macro*, es decir, un sistema macromolecular)**[104]**. Como ha observado Ernst Mayr [1960], hay tres razones diferentes para afirmar que una gran mutación es un gran paso en la librería de Mendel: produce una radical diferencia en el fenotipo (un monstruo); produce (por una u otra vía) un gran incremento de la capacidad adaptativa al medio; una gran cuantía de *elevación*, de acuerdo con nuestra metáfora para el *buen trabajo hecho* mediante cambios en el diseño.

Para la maquinaria de replicación molecular es posible dar grandes pasos en la librería de Mendel; hay casos en los que con un simple «error» de copia grandes fragmentos de texto cambian de lugar, invertidos o borrados. Es también posible que diferencias tipográficas se acumulen lentamente (y, en general, aleatoriamente) durante un largo período de tiempo en una gran porción del ADN que nunca se ha manifestado, y si estos cambios acumulados se expresan súbitamente, gracias a algún error de transposición, debemos esperar un enorme efecto en el fenotipo. Pero es solamente cuando retornamos al tercer sentido de la macromutación —es decir, grandes diferencias en la capacidad adaptativa— cuando se clarifica lo que parece ser

radical en la propuesta de Gould. Los términos «salto» y «macromutación» tienden a utilizarse para describir un movimiento seguido de éxito, un movimiento *creativo*, en el cual la progenie pasa, en una sola generación, de una región a otra del espacio de diseño, y *prospera como resultado*. La idea había sido promovida por Richard Goldschmidt [1933 y 1940] y se hizo inolvidable por su impactante expresión «monstruos prometedores». La notoriedad de este trabajo estriba en que Goldschmidt defendía la tesis de que tales saltos eran necesarios para que ocurriera la especiación.

Esta sugerencia fue rotundamente rechazada por la ortodoxia neodarwiniana, por las razones que ya hemos apuntado. Incluso antes de Darwin, la tradicional sabiduría de los biólogos sentenciaba, como escribió Linneo en su clásico trabajo de taxonomía (1751), que «*Natura non facit saltus*» —la naturaleza no da saltos— y esta era una máxima que Darwin no dejó intacta y proporcionó un enorme apoyo para ello. Los largos saltos laterales en un paisaje adaptativo casi nunca serán para nuestro beneficio; dondequiera que nos encontremos actualmente, estamos donde estamos debido a que ha sido una buena región en el espacio de diseño para nuestros ancestros —estamos cerca de la cima de uno u otro pico en este espacio—, así que mientras más grande sea el paso que demos (saltando de modo aleatorio, naturalmente), más probable es que, en cualquier caso, caigamos por un despeñadero hasta las tierras bajas (Dawkins 1986a:cap. 9). De acuerdo con este razonamiento estándar, no es un accidente que los monstruos no tengan, virtualmente, esperanzas. Esto es lo que hizo que los puntos de vista de Goldschmidt se consideraran heréticos; él conoció y aceptó que esto era verdad, en general, aunque propuso que, no obstante las extremadamente raras excepciones a esta regla, fueran los principales elevadores de la evolución.

Gould es un famoso defensor de desvalidos y de parias y ha deplorado el «ridículo ritual» (Gould 1982b:xv) a que fue sometido Goldschmidt por la ortodoxia darwiniana. ¿Trataba Gould de rehabilitar a Goldschmidt? Sí y no. En «Return of the Hopeful Monster», Gould se queja de que «los defensores de la teoría sintética de Darwin dibujan una caricatura de las ideas de Goldschmidt, escogido como víctima propiciatoria» (1980a:188). A muchos biólogos les pareció que Gould estaba argumentando que el equilibrio puntuado era una teoría sobre la especiación inspirada en Goldschmidt, a través de la macromutación. A estos biólogos les parecía que Gould trataba de agitar su admirable varita mágica de historiador sobre la empañada reputación de Goldschmidt para recuperar sus ideas en provecho propio. Aquí está el

mítico Gould, el refutador de la ortodoxia, marchando seriamente por el camino del real Gould, de modo que sus colegas sucumbían a la tentación de leer lo que había escrito con un grueso pincel. Sonreían con incredulidad y entonces, cuando Gould negó que estuviera apoyando la teoría del salto de Goldschmidt —en realidad, siempre la había apoyado— se mofaron irónicamente. Sabían lo que Gould había dicho.

Pero ¿realmente lo sabían? Debo admitir que pensé que habían leído sus escritos, hasta que Gould insistió en que revisara *todas* sus varias publicaciones, y comprobara por mí mismo que sus oponentes estaban imponiendo falsamente una caricatura de él. Me tocó en un punto débil; nadie conoce mejor que yo lo frustrante que resulta que los escépticos te cuelguen una burda pero oportuna etiqueta en uno de tus sutiles puntos de vista. (Yo soy ese tipo que tiene la reputación de negar que la gente tenga la experiencia de los colores o de los dolores y que creo que los termostatos piensan; ¡preguntadles a mis críticos!). Así que revisé los trabajos de Gould. «La ruptura de Goldschmidt» (Gould 1980b) fue el título que Gould eligió para esta renegación del gradualismo y recomendó, por serias consideraciones —sin apoyarlas—, algunas tesis radicales de Goldschmidt, aunque en el mismo artículo tuvo el cuidado de decir: «Nosotros no aceptamos todos sus argumentos acerca de la naturaleza de la variación». En 1982, Gould dijo claramente que el único aspecto de la tesis de Goldschmidt que él apoyaba era la idea de «pequeños cambios genéticos que producían grandes efectos por alterar el ritmo del desarrollo» (Gould 1981d:338), y en su introducción a la reimpresión del conocido libro de Goldschmidt, Gould se extiende en este punto:

> Los darwinistas, con su tradicional preferencia por el gradualismo y la continuidad, no pueden gritar hosannas ante la posibilidad de que se puedan producir en la evolución grandes desviaciones de fenotipos inducidas rápidamente por pequeños cambios genéticos que afectan al inicio del desarrollo; pero no hay nada en la teoría de Darwin que impida estos acontecimientos, ya que permanece la subyacente continuidad de los pequeños cambios genéticos (Gould 1982b:XIX).

En otras palabras, nada revolucionario:

> Pido disculpas por replicar: «Así que ¿qué es lo nuevo?». ¿Han negado este hecho algunos biólogos? Pero… el progreso en la ciencia exige, a menudo, la recuperación de antiguas verdades y su presentación bajo nuevas formas (Gould 1981d:343-344).

Por todo esto, Gould no puede resistir el impulso de describir este hecho, posiblemente infravalorado, acerca del desarrollo como una fuerza creativa no darwiniana en evolución, «porque las limitaciones que impone sobre la

naturaleza del cambio fenotípico garantizan que una variación darwiniana pequeña y continua no sea la materia prima de toda evolución», y por esto «relega la selección a un papel negativo (eliminación del incapaz) y asignan el mayor aspecto creativo de la evolución a la misma variación» (Gould 1982d:340).

Aún no queda claro cuál es la importancia que se le asigna a esta posibilidad, pero, en cualquier caso, Gould *no* ha proseguido por este camino: «El equilibrio puntuado no es una teoría de la macromutación» (Gould 1982c:88). Sin embargo, la confusión todavía abunda en esta valoración y Gould se ha visto obligado a seguir hablando de sus renuncias: «Nuestra teoría no implica un nuevo o violento mecanismo sino que sólo representa la ubicación a escala de los acontecimientos ordinarios en la vastedad del tiempo geológico» (Gould 1992b:12).

Así que todo esto ha sido una falsa alarma de revolución que en gran parte, si no en su totalidad, se encontraba ante los ojos de los espectadores. Pero en este caso, el punto de vista que hemos visto que mantenían Gould y Eldredge, una vez que hemos ampliado con el zoom, más allá de la falsa apariencia del tiempo compactado de sus diagramas biológicos, no es, después de todo, el escalón situado en el extremo derecho de la figura 10.8, sino uno de los otros caminos —gradual y nada violento— representado en la ilustración. Como ha señalado Dawkins, la forma en que Gould y Eldredge desafiaron al «gradualismo» no fue, en último término, a través de la presentación de un estimulante *no* gradualismo, sino diciendo que la evolución, cuando ocurre, era desde luego gradual, pero que la mayor parte del tiempo no era *ni siquiera* gradual; era una parada total. En la figura 10.6, el diagrama de la izquierda se supone que representa la ortodoxia darwiniana, pero la característica de esta ortodoxia que la teoría de Gould y Eldredge desafía no es el gradualismo, ya que, una vez que hemos obtenido la conversión del diagrama a la escala correcta, ellos mismos son también gradualistas. Con su teoría del equilibrio puntuado lo que Gould y Eldredge desafiaban era aquello que ya Dawkins (1986a:244) había denominado la «tesis de una velocidad constante» en el proceso de la evolución.

¿Ha estado comprometida la ortodoxia neodarwiniana con la tesis de una velocidad constante en el proceso evolutivo? En el artículo original, Eldredge y Gould sostenían que los paleontólogos se equivocaban al pensar que la ortodoxia requería de la tesis de una velocidad constante en la evolución. ¿Estaba el propio Darwin comprometido con la tesis de una velocidad constante? Con frecuencia Darwin insistía, correctamente, en su opinión de

que la evolución *tan sólo* puede ser gradual (en el mejor de casos, debería decirse). Como dice Dawkins: «Para Darwin, cualquier evolución que ha de ser ayudada por Dios para dar saltos no es, en absoluto, una evolución. Convertiría en un sin sentido el punto central de la evolución. A la luz de lo anterior, es fácil comprender por qué Darwin reiteraba constantemente la *gradualidad* de la evolución» (1986a: 145). Aunque no es precisamente difícil encontrar evidencia documentada en apoyo de la afirmación de que Darwin estaba comprometido con la tesis de una velocidad constante; hay un famoso pasaje en el cual Darwin expresa claramente el punto de vista opuesto, ese que puede ser llamado, en dos palabras, equilibrio puntuado:

Muchas especies una ve2 formadas nunca sufren cambio posterior alguno… y los períodos durante los cuales han sufrido modificación, aunque largos medidos en años, han sido probablemente cortos en comparación con los períodos durante los cuales mantienen la misma forma (*El origen de las especies*, 4.ª ed. y ss.; véase Peckham 1959:727).

Es una ironía, sin embargo, que Darwin, en su *Origen*, incluyera un diagrama donde aparecen rampas continuadamente ascendentes. Steven Stanley, otro exponente importante de la teoría del equilibrio puntuado, reproduce este diagrama en su libro (Stanley 1981:36) y hace una explícita inferencia en su leyenda.

Como consecuencia de tales afirmaciones existe hoy, sin duda, una tradición que imputa la tesis de una velocidad constante en la evolución tanto al propio Darwin como a la ortodoxia neodarwiniana. Por ejemplo, Colin Tudge, un buen periodista científico, al comentar las recientes opiniones de Elizabeth Vrba acerca del pulso de la evolución, señala sus presumibles implicaciones para la ortodoxia de las investigaciones actuales sobre la evolución de los impalas y de los leopardos:

El darwinismo tradicional predecía una modificación sostenida de los impalas durante tres millones de años, aun sin cambio climático, debido a que este animal todavía necesita correr más deprisa que los leopardos. Pero, de hecho, ni los impalas ni los leopardos han cambiado mucho. Son demasiado versátiles como para preocuparse de los cambios climáticos y la competición entre ellos y con su propia clase no aporta suficiente presión selectiva —como Darwin suponía— para provocarles alteraciones (Tudge 1993:35).

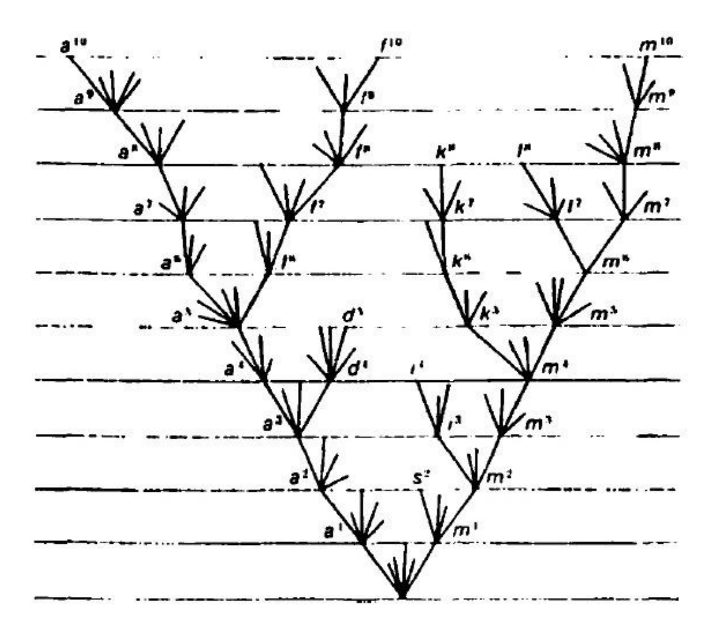


Figura 10.9

El árbol de la vida publicado por Darwin en *El origen de las especies*. (1859). El árbol despliega un modelo gradualista de la evolución. Cada disposición en abanico representa la lenta divergencia evolutiva de las poblaciones. Darwin creyó que nuevas especies, y eventualmente nuevos géneros y familias, se formaban mediante este tipo de divergencia lenta (Stanley 1981).

La presunción de Tudge de que el descubrimiento de tres millones de años de estasis en el impala y en el leopardo confundiría a Darwin está bastante extendida, aunque es un artefacto, directo o indirecto, de una lectura forzada de las «rampas» en los diagramas de Darwin y de otros ortodoxos.

Gould ha proclamado la muerte del gradualismo, pero ¿no es él mismo un gradualista, aunque no esté comprometido con la «tesis de una velocidad constante en la evolución»? Su negación de que su teoría proponga un «mecanismo violento», sugiere que sí es un gradualista, aunque sea duro decirlo, en la mismísima página en la que dice que, de acuerdo con la teoría del equilibrio puntuado,

> el cambio no ocurre habitualmente por alteración gradual imperceptible de especies enteras sino *más bien* [la cursiva es añadida] por el aislamiento de pequeñas poblaciones y su instantánea transformación biológica en nuevas especies (Gould 1992a:12).

Este pasaje de Gould nos invita a creer que el cambio evolutivo *no puede ser*, al mismo tiempo, «geológicamente instantáneo» y «gradualmente imperceptible». Pero esto es lo que el cambio evolutivo *debe ser* cuando no hay saltos. Dawkins dramatiza este punto con un asombroso experimento

mental del evolucionista G. Ledyard Stebbins, quien imagina un mamífero del tamaño de un ratón para el cual postula una presión de selección tan leve en favor de un aumento de su tamaño que no se produciría incremento mensurable por los biólogos que estudiaran el animal:

En lo que concierne a los científicos que estudiasen la evolución sobre el terreno, estos animales no estarían en absoluto evolucionando. Sin embargo, los animales evolucionan, a este ritmo lento, y llegarían a alcanzar eventualmente el tamaño de un elefante. ¿Cuánto tiempo tardarían?… Stebbins calculó que, a ese ritmo evolutivo muy lento, tardarían en alcanzar el tamaño del elefante el tiempo de unas 12.000 generaciones… Asumiendo para cada generación un período de cinco años, que es más largo que el del ratón aunque más corto que el del elefante, las 12.000 generaciones se extenderían durante 60.000 años. Estos 60.000 años suponen un período demasiado corto para ser medido por los métodos geológicos ordinarios, mediante la datación de fósiles. Como dice Stebbins: «El origen de un nuevo tipo de animal en un período de 100.000 años, o menos, es considerado por los paleontólogos como "súbito" o "instantáneo"» (Dawkins 1986a:242).

Parece cierto que Gould no calificaría como una violación del gradualismo este cambio localmente imperceptible del ratón en elefante, pero en este caso su propia oposición al gradualismo queda sin apoyo alguno en el registro de los fósiles. De hecho, Gould [1982a:383] admite esto ya que la única evidencia que su propio campo de la paleontología le proporciona, como oposición al gradualismo, va en dirección errónea. Gould puede desear disponer de evidencias acerca de una aceleración revolucionaria de un tipo o de otro, pero el registro de los fósiles sólo muestra períodos de estasis que sugieren que la evolución, a menudo, no es tan gradual como se piensa.

Pero quizá se le pueda dar un buen uso a este inconveniente: ¡quizás el desafío a la ortodoxia neodarwiniana no deba dirigirse a que no pueda explicar la puntuación, sino a que no pueda dar una explicación del equilibrio! Quizás el reto de Gould a la síntesis moderna del darwinismo deba ser que, después de todo, está comprometida con la tesis de una velocidad constante en el proceso evolutivo: que aunque Darwin no negó positivamente el equilibrio (aseguró que ocurría) no puede realmente explicar el equilibrio cuando éste tiene lugar, y tal equilibrio o estasis, puede argumentarse, es un modelo tan importante en el mundo que es necesario explicarlo. Esta es, de hecho, la nueva dirección en la que Gould orienta su ataque contra la síntesis moderna darwiniana.

¿Cómo podemos sostener que comprendemos la evolución si sólo estudiamos un uno o un dos por ciento de los fenómenos que construyen la historia direccional de la vida y dejamos el vasto campo de arbustos de la vida creciendo perfectamente —la historia de la mayoría de los linajes en la mayor parte del tiempo— en el limbo del olvido conceptual? (Gould 1993c:16).

Pero este camino tiene problemas. En primer lugar, hemos de evitar un error que es la imagen especular del error de Gould acerca del panadaptacionismo. No debemos caer en el error del «panequilibrismo». No obstante lo sorprendente o «invasivo» que resulte el modelo del estasis, sabemos por adelantado que la mayoría de los *linajes* no exhiben estasis. Están lejos de ello. Recordemos las dificultades con las que tropezamos cuando coloreamos a Lulu y a los de su misma especie en el capítulo 4. La mayoría de los linajes mueren pronto, sin tener nunca tiempo para establecer un período de estasis; sólo «vemos» una especie donde hay algo saliente y estable en el registro de los fósiles. El «descubrimiento» de que todas las especies exhiben estasis durante la mayor parte del tiempo es como el descubrimiento de que todas las sequías duran más de una semana. No tendríamos noticia de que *era* una sequía si no fuera un fenómeno de larga duración. De este modo, dado que una pizca de estasis es una precondición para la identificación de una especie, el hecho de que todas las especies exhiban algún grado de estasis es simplemente verdad por definición.

No obstante, el fenómeno del estasis puede ser un fenómeno real que necesite explicación. Podemos preguntarnos no por qué la especies exhiben estabilidad (algo verdadero por definición) sino por qué *hay* especies identificables y conspicuas, es decir; por qué los linajes consiguen la estabilidad. Pero, incluso en esta cuestión el neodarwinismo tiene varias y obvias explicaciones adaptacionistas para la pregunta de por qué el estasis debe ocurrir, a menudo, en un linaje. Lo primero lo hemos visto ya varias veces: toda especie es —debe ser— un negocio que marcha, y un negocio que marcha ha de ser conservador; la mayoría de las desviaciones, desde la tradición probada por el tiempo, serán rápidamente condenadas a la extinción. El mismo Eldredge [1989] ha sugerido que una importante razón para el estasis es el llamado *traslado de domicilio*. Sterelny lo describe así:

Cuando el medio ambiente cambia, los organismos pueden reaccionar trasladando su viejo hábitat. Pueden moverse hacia el norte cuando el clima refresca, más que evolucionando con adaptaciones al frío. (Esto no es un error: ¡Sterelny es un filósofo de la biología del Hemisferio Sur!). La selección dirigirá habitualmente el traslado. Los emigrantes que siguen el traslado de hábitat (personalmente o mediante dispersión reproductiva) estarán característicamente más adaptados que el fragmento de la población que no se mueve, ya que el fragmento residual estará menos bien adaptado al nuevo medio ambiente y se enfrentará con la nueva competición de otros emigrantes trasladando su viejo hábitat (1992:45).

Téngase en cuenta que el traslado del domicilio es una «estrategia» tanto de las plantas como de los animales. Naturalmente, algunos de los más claros casos de especiación invocan este fenómeno. Cuando el casquete polar se

encogió después de una edad del hielo, los límites de algunas plantas del norte de Asia se extendieron por el norte año tras año, «siguiendo» el hielo y extendiéndose al este y al oeste, cruzando la región del estrecho de Bering y quizá rodeando el globo como las gaviotas arenqueras. Entonces, mientras el hielo avanza hacia el sur durante la *próxima* edad de hielo, se rompen las conexiones entre las porciones de la familia situadas entre Asia y Norteamérica, creando dos hábitats aislados que naturalmente divergen en distintas especies, pero cuando ambos se mueven hacia el sur en sus respectivos hemisferios, continúan apareciendo los mismos, debido a que trasladan sus favorables condiciones climáticas, en lugar de permanecer con posteriores adaptaciones invernales[105].

Otra posible explicación del equilibrio puntuado es puramente teórica. Stuart Kauffman y sus colegas han desarrollado en el ordenador modelos que exhiben conductas en las cuales los períodos de estasis relativamente largos queden interrumpidos por breves períodos de cambio no desencadenados por ninguna interferencia del «exterior», así que este modelo parece ser una característica endógena o interna de la operación de este particular tipo de algoritmos evolutivos. (Para una reciente discusión de este tema, consúltese a Bak, Flyvbjerg y Sneppen 1994).

Queda completamente claro, en consecuencia, que para los neodarwinistas el equilibrio no representa ya un problema mayor que el que pueda presentar la puntuación; puede explicarse e incluso preverse. Sin embargo, Gould ha entrevisto otra revolución oculta en el equilibrio puntuado. Puede ser que los pasos horizontales de la puntuación no sean precisamente (relativamente) pasos rápidos en el espacio de diseño; puede que lo importante sea que se trate de pasos hacia la *especiación*. ¿Cómo podemos establecer una diferencia? Observemos la figura 10.10.

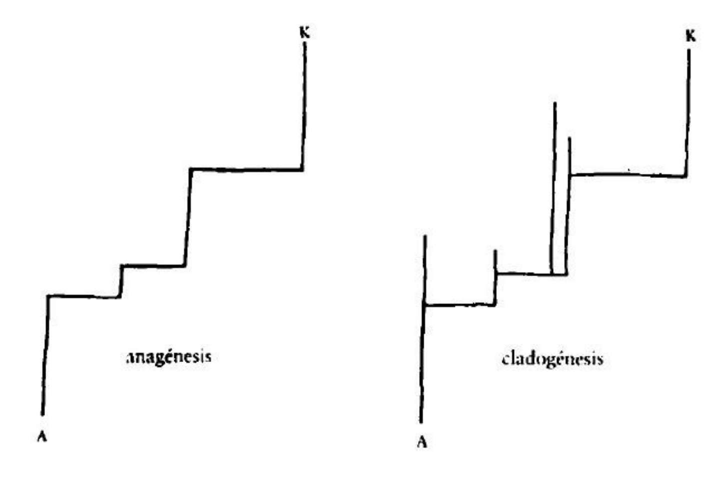


Figura 10.10

En ambos casos, el linaje que se dirige hacia *K* llega hasta donde llega siguiendo exactamente la misma secuencia de puntuaciones y equilibrios, pero el caso ilustrado a la izquierda muestra una *única especie* sometida a rápidos períodos de cambio, seguidos de largos períodos de estasis. Este tipo de cambio sin especiación se conoce como *anagénesis*. El caso ilustrado a la derecha de la figura es un ejemplo de *cladogénesis*, o sea, un cambio a través de la especiación. Gould sostiene que la tendencia hacia la derecha en los dos casos tendría una explicación diferente, Pero ¿cómo puede ser verdad? Recordemos lo que aprendimos en el capítulo 4: la especiación es un acontecimiento que sólo puede identificarse retrospectivamente. Nada de lo que sucede durante los *movimientos laterales* puede distinguir un proceso de anagénesis de uno de cladogénesis. Solamente ha habido especiación si existe un *posterior* crecimiento de ramas separadas que sobreviven lo suficiente para ser identificables como especies separadas.

¿No puede haber procesos especiales que pudieran ser denominados como especiación *prometedora* o especiación *incipiente*? Consideremos un caso en el que ocurra especiación. La especie paterna A se divide en las especies hijas B y C. Ahora, rebobinemos la cinta del proceso, justo lo suficiente en el tiempo para arrojar una bomba (un asteroide, un maremoto, una sequía, un veneno) sobre los primeros miembros de la especie B, como se representa en la zona media del diagrama. Al hacer esto convertimos lo que ha sido un caso de especiación en algo indistinguible de un caso de anagénesis (a la derecha, en la figura 10.11). El hecho de que la bomba impida ser abuelos a aquellos cuya progenie mata, difícilmente podría establecer una diferencia de cómo sus

contemporáneos llegan a separarse en grupos por presión selectiva. Esto requeriría algún tipo de causalidad retrotraída en el tiempo.

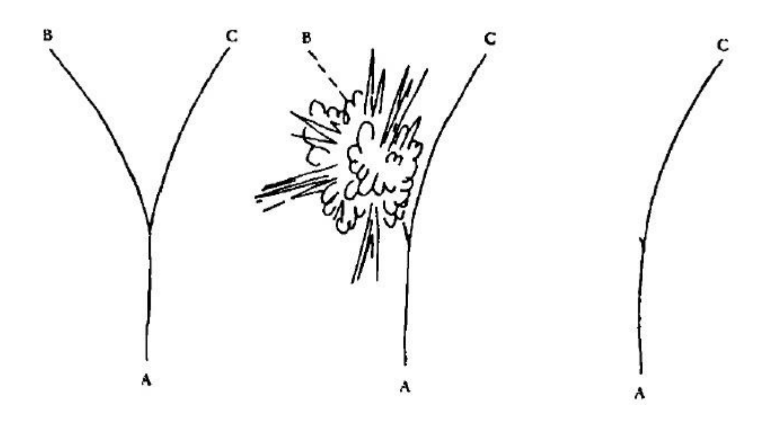


Figura 10.11

¿Es esto realmente verdad? Podemos creer que esto sería verdad si el acontecimiento que pone en marcha la especiación fuera una separación geográfica, capaz de garantizar el completo aislamiento causal de los dos grupos (especiación alopátrica), pero ¿qué pasaría si la especiación estuviera en marcha dentro de una población que formó dos subgrupos incomunicados desde el punto de vista reproductivo que compiten entre sí (en una forma de especiación simpátrica)? Darwin propuso, como ya señalamos (véase la página 58), que la competición entre formas muy relacionadas sería una fuerza impulsora en la especiación, así que la presencia —la no ausencia— de *lo que puede ser identificado retrospectivamente* como la primera generación de una especie «rival», puede ser muy importante, desde luego, para la especiación, pero el hecho de que estos rivales «vayan a ser» los fundadores de una nueva especie no puede desempeñar un papel en la intensidad o en otras características de la competición y, en consecuencia, en la velocidad o dirección del movimiento horizontal en el espacio de diseño.

Podemos suponer que los cambios morfológicos relativamente rápidos (movimientos laterales) son una precondición necesaria para la especiación. La rapidez del cambio se ve crucialmente afectada por el tamaño del fondo de genes; los grandes fondos de genes son conservadores y tienden a absorber los intentos de innovación sin dejar rastro. Un modo de reducir un gran fondo de genes es dividirlo en dos, y este puede ser el modo más frecuente de reducción de su tamaño, pero, de ahí en adelante, no se encuentra diferencia si la naturaleza descarta o no una de las dos mitades (como en el dibujo central de la figura 10.11). Es el cuello de botella de un disminuido fondo de genes el

que permite el movimiento rápido, no la presencia de dos o más cuellos de botella. Si hay especiación, entonces *dos especies completas* pasan a través de sus respectivos cuellos de botella; si no hay especiación entonces *una sola especie completa* es presionada contra un único cuello de botella. De modo que la cladogénesis, durante un período de puntuación, *no puede* entrañar un proceso diferente del que ocurre en la anagénesis, debido a que la diferencia entre cladogénesis y anagénesis es definible sólo en términos de *secuelas* tras la puntuación. Gould habla, a veces, como si la especiación marcase una diferencia. Por ejemplo, Gould y Eldredge [1993:225] se refieren al «requerimiento crucial de una supervivencia de los ancestros tras la ramificación puntuada» (como se muestra a la izquierda, en la figura 10.11), pero de acuerdo con Eldredge (comunicación personal) éste es tan sólo un requerimiento *epistemológico*, crucial para el teórico, ya que necesita la «supervivencia de los ancestros» como *evidencia* del descenso evolutivo.

Su explicación es interesante. El registro fósil está cargado de casos en los cuales una forma se detiene de modo brusco y otra forma, completamente diferente, aparece «en su lugar». ¿Cuáles de estos casos son saltos laterales rápidos en el proceso evolutivo y cuáles son casos de simple desplazamiento debido a la súbita inmigración de un pariente algo más distante? No podemos decirlo. Tan sólo cuando es posible observar lo que consideramos que son especies emparentadas coexistiendo durante algún tiempo con lo que creemos que es su progenie, podemos estar completamente seguros de que existe un camino directo desde la forma más precoz a la más tardía. Como cuestión epistemológica, ésta socava completamente la proposición que Gould hubiera deseado hacer: que la mayoría de los rápidos cambios evolutivos han tenido lugar a través de la especiación. Por esto, como afirma Eldredge, el registro fósil muestra *habitualmente* desviaciones bruscas *sin* ninguna «supervivencia de ancestros tras la ramificación puntuada» y si no hay datos acerca de cuáles son casos de cambio anagénico puntuado (como opuesto al fenómeno de la inmigración) entonces no hay modo de decir, a partir del registro de fósiles, si la especiación es un acompañamiento muy frecuente o muy raro de los cambios morfológicos rápidos**[106]**.

Aún pudiera existir otra vía que dé sentido a la insistencia de Gould de que es la especiación, y no la mera adaptación, lo que marca la gran diferencia en la evolución. ¿Qué pasaría si algunos linajes siguiesen un proceso de muchas puntuaciones (y en el proceso se produjesen muchas especies hijas) y otros linajes no, por lo cual tenderían a morir? Habitualmente los neodarwinistas asumen que las adaptaciones ocurren por la

transformación gradual de los organismos en linajes particulares, pero «si los linajes no cambian por transformación, entonces las tendencias a largo plazo en los linajes difícilmente pueden ser el resultado de su lenta transformación» (Sterelny 1992:48). Esta ha sido considerada una posibilidad interesante (en su artículo original, Eldredge y Gould la discuten muy brevemente, y dan crédito a Sewall Wright 1967 como una de sus fuentes). La versión de Gould [1982a] de esta idea es que la especie completa no se revisa por un rediseño, pieza por pieza, de sus miembros individuales; las especies son cosas inmodificables y bastante quebradizas; las desviaciones en el espacio de diseño suceden (¿ampliamente?, ¿a menudo?, ¿siempre?) a través de la *extinción* y el *nacimiento* de las especies. Esta hipótesis es la que Gould y Eldredge [1993:214] llaman «separación a un nivel más alto». A veces se le denomina selección de especies. Es difícil clarificar este punto central, pero ya tenemos a mano el equipamiento para intentarlo. Recordemos lo ya dicho acerca del papel en la evolución de la «táctica del palo y la zanahoria». En efecto, Gould está proponiendo una nueva aplicación de esta idea fundamental de Darwin; no creamos que la evolución *hace ajustes* sobre linajes existentes; la evolución *elimina* linajes completos y permite prosperar a otros. Parece como si hubiera ajustamientos en linajes a lo largo del tiempo, pero lo que realmente está funcionando es el sistema del «palo y la zanahoria» a nivel de las especies. El *nivel correcto* desde el cual observar la tendencia evolucionaría, puede entonces Gould afirmar, no es el nivel del gen o del organismo, sino el nivel de las especies completas o «clades». En lugar de dirigir la atención a la pérdida de genes concretos en los fondos de genes, o a la muerte diferencial de genotipos concretos dentro de una población, se trataría de observar el ritmo de la extinción diferencial de especies completas y el «nacimiento» diferencial de especies; el ritmo con el cual un linaje pueda originar, por especiación, especies hijas.

Esta es una idea interesante, pero, a primera vista, no es un negación de la afirmación de la tesis ortodoxa que defiende que las especies sufren una transformación por la vía del «gradualismo filético». Puede ser verdad, como Gould propone, que algunos linajes generen muchas especies hijas y otros no, y que los primeros tiendan a sobrevivir más que los segundos. Observemos la trayectoria, a través del espacio de diseño, de cada especie superviviente. La especie completa se encuentra en cualquier período de tiempo en estasis o sometida a un cambio puntuado, pero este cambio, después de todo, es, en sí mismo, «una lenta transformación del linaje». Puede ser verdad que la mejor vía para contemplar el patrón macroevolucionario a largo plazo sea la de

comprobar las diferencias en la «fecundidad del linaje», en lugar de observar la transformación de los linajes individuales. Esta es una propuesta poderosa que merece tomarse en serio, aunque no refuta ni suplanta el gradualismo; lo construye[107].

(El desplazamiento de nivel propuesto por Gould me recuerda el desplazamiento de nivel entre *hardware* y *software* en la ciencia informática; el nivel del *software* es el nivel correcto desde el cual se responde a ciertas cuestiones a gran escala, pero no arroja duda alguna sobre la verdad de la explicaciones de los mismos fenómenos a nivel del *hardware*. Estaríamos locos si tratásemos de explicar las diferencias visibles entre el WordPerfect y el Microsoft Word a nivel del *hardware* y quizá yo estaría loco si tratara de explicar algunos de los modelos visibles de la diversidad de la biosfera concentrándome en la lenta transformación de varios linajes, pero esto no significa que no sufriesen lentas transformaciones a lo largo de su historia, en varias marcas de puntuación).

La relativa importancia de una selección de las especies del tipo que Gould está ahora proponiendo aún no ha sido determinada. Y está claro que, a pesar del amplio papel que la selección de las especies desempeñará en las últimas versiones del neodarwinismo, no se trata de un gancho celeste. Después de todo, el modo según el cual los nuevos linajes entran en escena como candidatos para la selección de especies es por la micromutación gradualista estándar, a menos que Gould desee abrazar a monstruos prometedores. De este modo, Gould puede haber contribuido a descubrir una nueva grúa, si esto es lo que resulta ser: un mecanismo de innovación de diseño, hasta ahora inapreciado y no reconocido, construido a partir de mecanismos ortodoxos y de tipo estándar. Sin embargo, dado que mi diagnóstico es que Gould ha estado todo el tiempo esperando encontrar ganchos celestes y no grúas, cabe predecir que continuará buscando. ¿Puede haber, quizás, algo más acerca de la especiación que sea tan especial que no pueda ser manejado por el neodarwinismo? La explicación de la especiación por Darwin, como ya he recordado, invoca la competición entre parientes cercanos.

Las nuevas especies habitualmente ganan un domicilio expulsando a otras en abierta competición (un proceso que Darwin a menudo describió en sus libros de notas como «metiendo una cuña»). Esta constante batalla y conquista aporta una base racional para el progreso, dado que los vencedores, por término medio, pueden asegurar sus éxitos gracias a la superioridad general de su diseño (Gould 1989b:8).

A Gould no le gusta esta imagen de «meter una cuña» en las otras especies. ¿Qué hay de malo en ella? Bien, es una invitación —afirma Gould— a la creencia en el progreso, pero esta invitación, como hemos ya visto, es tan fácilmente rechazada por el neodarwinismo como lo fue por el mismo Darwin. El progreso global a largo plazo, es decir, el punto de vista de que las cosas en la biosfera, por lo general, van cada vez mejor, fue negado por Darwin y aunque, a menudo, los espectadores lo imaginan como una implicación de la teoría de la evolución, se trata sencillamente de un error, en el cual cae el darwinismo no ortodoxo. ¿Qué puede haber de erróneo en la imagen de la cuña? En el mismo artículo Gould habla del «trabajo laborioso y previsible de la cuña», y yo sugiero que precisamente es esto lo que le ofende de esta imagen: al igual que la rampa del gradualismo, esta imagen sugiere una suerte de marcha laboriosa, predecible y no inteligente, por las laderas del espacio de diseño (véase, por ejemplo, Gould 1993d:cap. 21). El problema con una cuña es sencillo: no se trata de un gancho celeste.

4. *Tinker pasa a Evers y Evers pasa a Chance: el misterio del doble juego en Burgess Shale*[108]

> Aún hoy, un buen número de mentes distinguidas parecen incapaces de aceptar e incluso comprender que, a partir de una fuente ruidosa, la selección natural, sola y sin ayuda, haya podido extraer toda la música de la biosfera. De hecho, la selección natural actúa sobre productos del azar y no puede alimentarse en cualquier parte; aunque opera en un terreno de condiciones muy exigentes, donde el azar está prohibido.
>
> Jacques Monod,
> *El azar y la necesidad*

> Pero el puntualismo moderno —especialmente cuando se aplica a los caprichos de la historia humana— resalta el concepto de contingencia: la imposible predicción de la naturaleza de la estabilidad futura, y el poder de los acontecimientos contemporáneos y las características personales para conformar y dirigir el camino real seguido entre miles y miles de posibilidades.
>
> Stephen Jay Gould,
> «*Life in a Punctuation*»

En esta cita, Gould no se refiere precisamente a la imposible predicción sino al poder de los acontecimientos contemporáneos y las *características individuales* para «dar formar y dirigir el camino real» de la evolución. Todo esto suena exactamente como el eco de la esperanza que condujo a James

Mark Baldwin a descubrir el efecto que ahora lleva su nombre: *de algún modo* hemos de aceptar que las *características individuales* —consciencia, inteligencia, acción— dirigen el proceso evolutivo, aunque desde la grada del estadio. ¡Si podemos tener contingencia —una contingencia radical—, ésta proporcionará a la *mente* algún espacio para que pueda *actuar* responsabilizándose de su propio destino, en lugar de aparecer como el mero efecto de una cascada no inteligente de procesos mecánicos! Esta conclusión, me atrevo a sugerirlo, es el objetivo último de Gould, revelado en los caminos que más recientemente ha explorado.

En el capítulo 2 mencioné que la principal conclusión del libro de Gould titulado *La vida maravillosa* (1989) es que si rebobinamos la cinta de la vida pasándola una y otra vez, las posibilidades de que volviéramos a aparecer de nuevo son enormemente escasas. Hay tres cosas acerca de esta conclusión que han desconcertado a quienes la han revisado. Primero, ¿por qué cree Gould que esta conclusión es tan importante? (De acuerdo con la sobrecubierta: «En esta obra maestra, Gould explica por qué la diversidad del yacimiento de fósiles de Burgess Shale es importante para la comprensión de la cinta de nuestro pasado y para la remodelación del camino por el que ponderamos la criba de la existencia y la pasmosa improbabilidad de la evolución humana»). En segundo lugar, ¿cuál es exactamente su conclusión y a quiénes se refiere Gould cuando dice «nosotros»? Y, en tercer lugar, ¿cómo piensa Gould que esta conclusión se deriva (cualquiera que sea) de su fascinante discusión sobre el yacimiento de fósiles de Burgess Shale, con la que no parece tener relación alguna? Analizaremos estas cuestiones, a nuestro modo, desde la tercera base a la segunda y desde ésta a la primera[109].

Gracias al libro de Gould, Burgess Shale, una cantera de pizarra en la falda de una montaña en la Columbia británica, se ha convertido de lugar famoso para los paleontólogos en un santuario internacional de la ciencia, el lugar del nacimiento de… ¡bien!, de Algo Realmente Importante. Los fósiles allí encontrados datan del período conocido como la explosión cámbrica, un período que tuvo lugar hace alrededor de seiscientos millones de años, cuando los organismos multicelulares despegaron realmente, creando las ramas palmeadas del árbol de la vida en la figura 4.1. Formados bajo condiciones peculiarmente afortunadas, los fósiles inmortalizados en Burgess Shale son bastante más completos y tridimensionales de lo que son habitualmente, y la clasificación que de ellos hizo Charles Walcott en este siglo se basó en la disección literal de algunos de estos fósiles. Walcott introdujo muy forzadamente a las variedades que encontró en los *phyla* tradicionales y así

quedó la cosa —es decir, toscamente clasificados— hasta las brillantes reinterpretaciones realizadas, en las décadas de los setenta y los ochenta, por Harry Whittington, Derek Briggs y Simon Conway Morris, quienes afirmaron que muchas de estas criaturas —que constituyen un lote impresionantemente extraño y extravagante— habían sido erróneamente clasificadas; en realidad pertenecen a *phyla* que no tienen descendencia moderna, y no habían sido imaginados anteriormente.

Todo esto es fascinante, pero ¿qué hay en ello de revolucionario? Gould ciertamente pensó que lo era: «Creo que la reconstrucción de la *Opabinia* realizada por Whittington en 1975 es uno de los grandes documentos en la historia del conocimiento humano» (1989a:136). Su trío de héroes no estaba en esta línea (véase, por ejemplo, Conway Morris 1989) y la precaución que mostraron se ha revelado profética; análisis posteriores han atemperado, después de todo, algunas de las más radicales hipótesis reclasificatorias (Briggs y otros 1989, Foote 1992, Gee 1992, Conway Morris 1992). Si no fuera por el pedestal en el que Gould había colocado a sus héroes, no darían ahora la impresión de haber caído desde tan alto; el primer escalón era de calidad inmejorable y no fueron capaces de alcanzarlo por sí mismos.

Pero, en cualquier caso, ¿cuál era la cuestión revolucionaria que Gould pensó que quedaba establecida gracias a lo que habíamos aprendido acerca de estas criaturas cámbricas? La fauna de Burgess Shale apareció súbitamente (recuérdese lo que esto significa para un geólogo) y la mayor parte desapareció también súbitamente. Esta entrada y salida no gradual, sostiene Gould, demuestra la falacia de lo que él llama «el cono de la creciente diversidad», ilustrando esta tesis con un notable par de árboles de la vida.

Una figura vale por mil palabras, y Gould subraya, una y otra vez, con muchas ilustraciones, el poder de la iconografía para engañar incluso al experto. La figura 10.12 es otro ejemplo y este es suyo. En la parte superior, nos dice Gould, se encuentra el viejo y falso punto de vista, el cono de la creciente diversidad; en la parte baja, el punto de vista mejorado de la diezma y la diversificación. Pero nótese que podemos convertir el dibujo de la parte inferior en un cono de diversidad creciente simplemente *alargando la escala vertical*. (Alternativamente, podemos transformar el dibujo de la parte superior en un nuevo y aprobado icono de la inferior si comprimimos hacia abajo la escala vertical, al estilo de los diagramas tipo estándar de equilibrio puntuado, como en el ejemplo de la derecha en la figura 10.6.) Dado que la escala vertical es arbitraria, los diagramas de Gould no ilustran, en absoluto, ninguna diferencia. La mitad inferior de su diagrama inferior ilustra

perfectamente un cono de diversidad creciente, y ¿quién sabe si la fase siguiente de actividad en el diagrama superior sería una diezma que se convierte en una réplica del diagrama inferior?

El cono de diversidad creciente no es obviamente una falacia, si medimos la diversidad por el número de especies. Antes de que hubiera cien especies hubo diez, y antes de que hubiera diez hubo dos, y así debió ser en cada rama del árbol de la vida. Las especies se van extinguiendo continuamente y quizás el 99 por ciento de todas las que han existido alguna vez estén ahora extinguidas, por lo que debemos contar con muchas diezmas para equilibrar la diversificación. El florecimiento y la desaparición del yacimiento de fósiles de Burgess Shale puede haber sido menos gradual que el de otras faunas, antes o después, pero esto no muestra nada radical acerca de la forma del árbol de la vida.

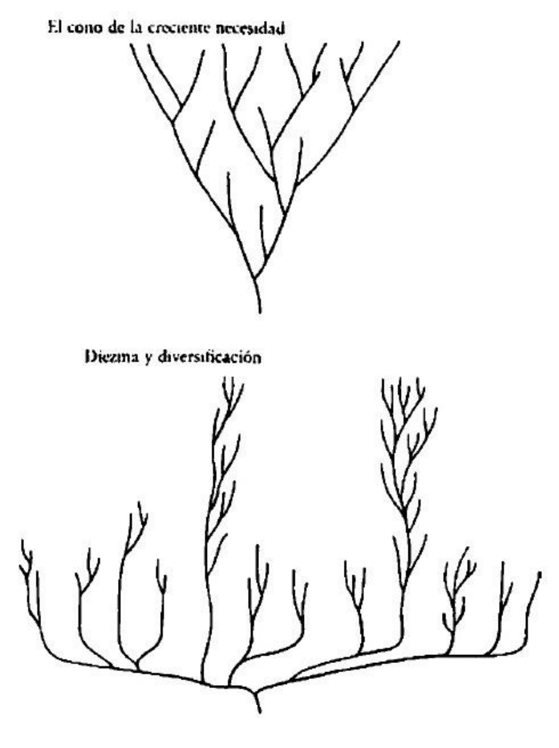


Figura 10.12

El cono formado por la creciente diversidad. La falsa, aunque convencional iconografía del cono de la creciente diversidad, y el modelo de di versificación y diezma, sugerido por la correcta reconstrucción a partir de los hallazgos de Burger Shale. (Gould 1989a:46).

Alguien pudiera decir que esta opinión no tiene en cuenta el siguiente punto de Gould: «¡Lo especial de la espectacular diversidad de la fauna de Burgess Shale es que no se trataba precisamente de nuevas especies sino de

phyla totalmente nuevas! ¡Eran *radicalmente* nuevos diseños!». Confío en que éste no fuera nunca el punto de vista de Gould, porque si lo fuera se trataría de una embarazosa falacia de coronación retrospectiva; como ya hemos visto, todos los nuevos *phyla* —¡naturalmente, nuevos reinos!— han de comenzar como meras subvariedades para, a partir de ahí, llegar a ser nuevas especies. El hecho de que, desde la ventaja de los puntos de vista actuales, aparezcan como miembros precoces de nuevos *phyla*, no tiene, en absoluto, nada de especial. Pueden ser *especiales*, sin embargo, no porque estuvieran destinados a ser los fundadores de nuevos *phyla*, sino porque eran, sorprendentemente, *morfológicamente* diversos. Según Gould, el modo de probar esta hipótesis sería, como Dawkins [1990] había dicho, «aplicar la regla sobre los mismos animales, sin prejuicios por las modernas preconcepciones acerca de los "fundamentales planos corporales" y su clasificación. El verdadero índice de hasta qué punto dos animales no son similares es comprobar ¡en qué medida no son realmente similares!». Sin embargo, los estudios realizados hasta la fecha sugieren, que, de hecho, la fauna de Burgess Shale, con todas sus peculiaridades, no exhibe una inexplicable o revolucionaria diversidad morfológica (por ejemplo, Conway Morris 1992, Gee 1992, McShea 1993).

La fauna de Burgess Shale fue azotada, permítasenos suponerlo aunque no se sepa realmente, por una de las periódicas extinciones en masa que han tenido lugar en la Tierra. Los dinosaurios, como todos sabemos, sucumbieron más tarde, en la Extinción del Período Cretáceo, (conocida también como la extinción en los límites K-T), probablemente provocada, hace alrededor de 65 millones de años, por el impacto de un enorme asteroide. La extinción masiva de Burgess Shale le pareció a Gould muy importante y un reto para el neodarwinismo. «Si el equilibrio puntuado sienta mal a las expectativas tradicionales (¡y les sigue sentando mal ahora!) la extinción en masa les sienta peor» (Gould 1985:242). ¿Por qué? Según Gould, la ortodoxia requiere «extrapolacionismo», la doctrina de que todo cambio evolutivo es gradual y predecible: «Pero si las extinciones en masa son verdaderas rupturas de la continuidad, si la lenta construcción de la adaptación en tiempos normales no se extiende con los previstos éxitos a través de los límites de la extinción en masa, entonces el extrapolacionismo falla y el adaptacionismo sucumbe» (Gould 1992a:53). Esto es totalmente falso, como ya he señalado:

No puedo comprender cómo un adaptacionista puede incurrir en la locura de apoyar algo como el «extrapolacionismo» en una forma tan «pura» como para negar la posibilidad, o incluso la probabilidad, de que la extinción en masa podría tener un papel importante en el crecimiento del árbol de la vida, como señala Gould. Ha sido siempre obvio que el dinosaurio más perfecto sucumbiría si un cometa se

estrellara en la tierra donde vive con una fuerza cientos de veces superior a la de todas las bombas de hidrógeno fabricadas hasta ahora (Dennett 1993b:43).

Gould [1993e] respondió citando un pasaje de Darwin que expresa con claridad puntos de vista extrapolacionistas. ¿Está hoy el adaptacionismo comprometido con esta implicación desesperanzada? He aquí un ejemplo de cómo el propio Charles Darwin ha de contar como un hombre de paja, ahora que los neodarwinistas se han movido. Es verdad que Darwin tenía tendencia a insistir, falto de visión a largo plazo, en la naturaleza gradual de toda extinción, pero ha sido reconocido ampliamente por los neodarwinistas que esta actitud se debía a su vehemencia por separar sus puntos de vista de las variantes del catastrofismo, que son un obstáculo para aceptar la teoría de la evolución por selección natural. Debemos recordar que en los días de Darwin los milagros y las calamidades, tales como el Diluvio bíblico, eran los principales rivales del pensamiento darwiniano. No es de extrañar que Darwin tendiese a rechazar cualquier cosa que pareciese sospechosamente rápida y conveniente.

El hecho —si fue así— de que la fauna de Burgess Shale hubiera sido diezmada en una extinción masiva es, en cualquier caso, menos importante para Gould que otra conclusión que él desea extraer acerca de este hecho: fue diezmada, sostiene (1989a:47n), aleatoriamente. De acuerdo con el punto de vista ortodoxo, «los supervivientes ganaron» aunque Gould opina: «Quizá la Parca que con su guadaña liquidaba los diseños anatómicos fuera solamente la diosa de la Fortuna disfrazada» (1989a:48). ¿Fue, en verdad, este episodio una lotería que determinó todas sus suertes? Esta sería una tesis maravillosa y definitivamente revolucionaria, especialmente si Gould la convirtiese en una generalización, pero la realidad es que no posee evidencia alguna para una tesis tan fuerte y retrocede:

Estoy dispuesto a dar por hecho que algunos grupos pueden haber disfrutado de una ventaja (aunque no tengo idea de cómo identificarla o definirla), pero sospecho que [la hipótesis de la diosa de la Fortuna] encierra una verdad central acerca de la evolución. Burgess Shale al hacer esta… interpretación inteligible por el hipotético experimento de la cinta, promueve una visión radical de las vías de la evolución y de la previsibilidad (1989a:50).

La sugerencia de Gould, por lo tanto, no es que él pueda probar la hipótesis de la diosa de la Fortuna, sino que los hallazgos de Burgess Shale la hacen, al menos, inteligible. Sin embargo, tal como Darwin insistió desde el principio, lo único que se necesita son «algunos grupos» con una «ventaja» para poner en marcha la cuña de la competición. ¿Está diciendo Gould que la mayor

parte de la competición (o la competición con sus mayores y más importantes efectos) fue una auténtica lotería? Esto es lo que «sospecha».

¿Cuál es su evidencia para apoyar esta sospecha? Absolutamente ninguna. Lo que Gould ofrece es el hecho de que él, observando esas asombrosas criaturas, no puede imaginar por qué algunas habrían de estar mejor diseñadas que otras. Todas le parecen igualmente extrañas y desgarbadas. Dados sus respectivos predicamentos, esta no es una buena evidencia de que no difieran claramente, de hecho, por la calidad de su ingeniería. Si no estamos dispuestos a realizar ingeniería revertida, no estaremos en condiciones de concluir que *no* hay explicación de ingeniería revertida que pueda demostrarse. Gould ofrece una apuesta: «Reto a cualquier paleontólogo a argumentar que él podría volver atrás en el tiempo a los mares que cubrían Burgess Shale y, sin el beneficio de una visión retrospectiva, escoger *Naroia, Canadaspis, Aysheaia y Sanctaris* como las especies que tendrían éxito, mientras que identificaría a *Maarrella, Odaraia, Sidneyia y Leonchoilia* como especies maduras para ser segadas por la guadaña de la Parca» (1989a:188). Esta es una apuesta apetitosa, muy segura, dado que los paleontólogos lo que deberían examinar serían los contornos de los órganos visibles en las huellas fósiles. Pero estos contornos podrían haberse perdido. Algunos ingenieros realmente ingeniosos en ingeniería revertida algún día serán capaces de contarnos una historia excesivamente convincente acerca de por qué los ganadores ganaron y los perdedores perdieron. ¿Quién sabe? Todos sabemos una cosa: no podemos hacer una revolución científica a partir de una corazonada que casi no puede someterse a prueba. (Véanse también Gould 1989a:238-239, y Dawkins 1990 para más observaciones sobre este tema).

De modo que seguimos tropezando con un misterio acerca de lo que Gould piensa que es tan especial en el único florecimiento y desaparición de esas asombrosas criaturas, las cuales le inspiran una sospecha, pero ¿por qué? He aquí una clave, extraída de una charla que Gould dio en el Edinburg International Festival of Science and Tecnology, titulada «El individuo en el mundo de Darwin»:

e hecho casi todos los diseños anatómicos más importantes de organismos aparecen en una gran exhalación denominada la Explosión Cámbrica, que tuvo lugar hace unos 600 millones de años. Debemos tener en cuenta que una exhalación o una explosión en términos geológicos necesita una mecha muy larga. Pudo tratarse de unos dos millones de años, aunque dos millones de años comparados con la línea evolutiva de miles de millones es nada. *Y no es a esto a lo que este mundo de avance necesario y predecible debería asemejarse* (Gould 1990:12). (La cursiva es mía).

¿Es así, realmente? Consideremos una situación paralela. Nos sentamos en una roca en el Yellowstone National Park, en Wyoming, observando un agujero en el suelo. Nada sucederá durante diez, veinte, treinta minutos, y entonces, súbitamente —¡una exhalación!—, un chorro de agua hirviendo se dispara a más de treinta metros de altura. Todo ha terminado en unos pocos segundos y luego no sucede nada más —aparentemente estamos como antes — y esperamos durante una hora, pero nada sucede. Esta fue, entonces, nuestra experiencia: una única y extraordinaria explosión que duró pocos segundos en el transcurso de una hora y media de aburrimiento. Quizá tendríamos la tentación de pensar: «¡Seguramente éste ha sido un acontecimiento único e irrepetible!».

Entonces ¿por qué llaman el Old Faithful (‘el Viejo Fiel’) a este géiser de Wyoming? De hecho, este géiser se repite a sí mismo, año tras año, cada sesenta y cinco minutos por término medio. La «forma» de la Explosión Cámbrica —su «súbito» comienzo y su también «súbito» final— no es evidencia, *en absoluto*, de la tesis de la «contingencia radical». Pero Gould parece creer que sí lo es**[110]**. Gould parece creer que, si volviéramos a pasar la cinta de la vida, no podremos conseguir otra Explosión Cámbrica la próxima vez, o nunca más. Pero aunque esto pudiera ser cierto, Gould todavía no ha presentado ni un solo *bit* de evidencia.

¿De dónde puede llegar tal evidencia? De las simulaciones de la vida artificial en el ordenador, por ejemplo, las cuales nos permiten rebobinar la cinta de la vida, una y otra vez. Es sorprendente que Gould haya pasado por alto la posibilidad de que puede encontrar alguna evidencia a favor (o en contra) de su principal conclusión, si observa el campo de la vida artificial, pero nunca mencionó esta posibilidad. ¿Por qué no lo hizo? No lo sé, aunque si sé que Gould no está muy encariñado con los ordenadores y que, hasta ahora, no ha utilizado el ordenador para el procesamiento de textos; este detalle pudiera tener algo que ver con su actitud.

Una clave mucho más importante, seguramente, es el hecho de que cuando hacemos retroceder la cinta de la vida, encontramos todo tipo de evidencias de repetición. Ya hemos visto que esto se debe a que la evolución convergente es la propia vía por medio de la cual la naturaleza vuelve a pasar la cinta. Como ha escrito Maynard Smith:

En el experimento de Gould de «volver a pasar la cinta del período cámbrico» mi predicción sería que muchos animales evolucionarían sus ojos, debido a que los mismos han evolucionado muchas veces, de hecho, en muchas clases de animales. Apostaría a favor de que algunos evolucionarían en la capacidad de volar, debido a que el vuelo ha evolucionado cuatro veces, en dos *phyla* diferentes; pero no tendría la

certeza debido a que los animales no pueden nunca salir de la tierra. Sin embargo, estoy de acuerdo con Gould que no se puede predecir qué *phyla* sobrevivirá y heredará la tierra (Maynard Smith 1992:34).

El último punto de Maynard Smith es artificioso: si reina la evolución convergente, ¡poco importa qué *phyla* herede la tierra, debido a la aplicación en el proceso evolutivo del «palo y la zanahoria»! Combinando el «palo y la zanahoria» con la evolución convergente llegaremos a la conclusión ortodoxa de que *cualquier* linaje que sobreviva gravitará hacia los buenos movimientos en el espacio de diseño, y será difícil decir el resultado a partir del que salga ganador allí, si algún linaje diferente ha seguido adelante. Consideremos, por ejemplo, el kiwi, un pájaro que ha evolucionado en Nueva Zelanda, donde no existen mamíferos que compitan con él, por lo que ha podido converger en un extraordinario número de características de los mamíferos; básicamente es un pájaro que pretende ser un mamífero. El mismo Gould (1991b) ha escrito acerca del kiwi y su huevo extraordinariamente grande, pero como Conway Morris puntualiza en su revisión:

Hay algo más acerca del kiwi que sólo se menciona de pasada, y es la extraordinaria convergencia entre los kiwis y los mamíferos… Estoy seguro de que Gould sería el último en negar la convergencia, aunque sin duda esto socava bastante su tesis de la contingencia (Conway Mollis 1991:6).

Gould no niega la convergencia —¿cómo iba a hacerlo?— pero tiende a pasarla por alto. ¿Por qué? Quizá porque, como dice Conway Morris, esta es la fatal debilidad de su defensa de la contingencia (Véanse también Maynard Smith 1992, Dawkins 1990, Bickerton 1993).

De modo que ahora tenemos una respuesta para nuestra tercera pregunta. La fauna de Burgess Shale inspira a Gould porque ha pensado erróneamente que aporta evidencia para su tesis de la «contingencia radical». *Podría* ilustrar la tesis pero esto no lo sabremos hasta que hagamos el tipo de investigación que el propio Gould ha dejado de lado. Hemos alcanzado la segunda base. Concretamente, ¿cuál es la propuesta de Gould acerca de la contingencia? Gould dice que «el malentendido más frecuente acerca de la evolución, al menos en la cultura de los legos», es la idea de que «nuestra eventual aparición» es «algo intrínsecamente inevitable y predecible dentro de los confines de la teoría» (1990:3). ¿*Nuestra* aparición? ¿Qué significa esto? Hay una escala deslizante sobre la cual Gould olvidó localizar su llamada acerca del rebobinado de la cinta. Si con «nosotros» quiere decir algo muy particular —Steve Gould y Dan Dennett, por ejemplo— entonces no necesitamos la hipótesis de la extinción en masa para convencernos de lo afortunados que somos por estar vivos; si nuestras dos madres nunca hubieran encontrado a

nuestros padres, eso bastaría para relegarnos a Neverland (la Tierra que Nunca Será) y, por supuesto, lo mismo sucedería con todos los seres humanos actualmente vivos. Si tan triste fortuna hubiera caído sobre nosotros, esto no significaría, sin embargo, que nuestros respectivos despachos en Harvard y en Tufts estarían vacantes. Sería asombroso si el nombre del ocupante del despacho de Harvard en esta circunstancia fuera «Gould» y yo no apostaría a favor de que su ocupante fuera un habitual de las boleras en Fenway Park, pero sí apostaría a que su ocupante sabe mucho de paleontología, que imparte conferencias y publica artículos y pasa miles de horas estudiando la fauna (la flora no; el despacho de Gould está en el Museo de Zoología Comparada). En cambio, si lo que Gould quiere decir con «nosotros» es algo muy general, algo así como un «vertebrado terrestre que respira», probablemente estaría equivocado, por las razones que Maynard Smith ha mencionado. De modo que suponemos que lo que Gould quiere decir es algo intermedio, algo así como «seres inteligentes, que usan lenguaje, inventan tecnología y crean cultura». Esta es una interesante hipótesis. Si esto es verdad, entonces, al contrario de lo que muchos pensadores suponen habitualmente, la búsqueda de inteligencia extraterrestre es tan quijotesca como la búsqueda de canguros extraterrestres; esto sucedió una vez, aquí, pero probablemente nunca más volverá a suceder. Pero el libro de Gould *La vida maravillosa* no ofrece evidencia a su favor (Wright 1990); aunque la diezma de la fauna de Burgess Shale hubiera sido aleatoria, cualesquiera que hubiesen sido los linajes supervivientes, según la teoría estándar neodarwinista, hubieran procedido a agruparse hacia los dotados con buenas mañas en el espacio de diseño.

Hemos respondido a nuestra segunda pregunta. Estamos listos, por fin, para atajar la primera base: ¿por qué sería esta tesis de tan gran importancia, cualquiera que sea la vía por la que se produzca? Gould cree que la hipótesis de la «contingencia radical» trastorna nuestra ecuanimidad, pero ¿por qué?

Hablamos acerca de «la marcha de la mónada hacia el hombre» (de nuevo el lenguaje de viejo estilo) como si la evolución hubiera seguido caminos de progreso continuado a lo largo de linajes intactos. Nada puede estar más lejos de la realidad (Gould 1989b:14).

¿*Qué* es lo que no puede estar más lejos de la realidad? A primera vista podría parecer como si Gould estuviera diciendo que no hay continuidad ni linaje intacto entre las «mónadas» y nosotros, pero seguramente la hay. No hay una implicación más segura que ésta de la gran idea de Darwin. Como subrayamos en el capítulo 8, no es motivo de controversia que todos somos directos descendientes de macros —o mónadas— simple replicadores

precelulares, bajo un nombre u otro. Por lo tanto, ¿qué es lo que está diciendo Gould? Quizá tengamos la intención de hacer hincapié sobre «caminos de *progreso*»; esta es la creencia en el progreso que está tan lejos de la verdad. Los caminos son continuos, los linajes completamente intactos, aunque no sean linajes de progreso global. Esto es verdad, pero ¿y qué?

No hay caminos *globales* de progreso, aunque sí hay un incesante progreso *local*. Este mejoramiento busca el mejor diseño con tan gran fiabilidad que el razonamiento adaptacionista puede, a menudo, predecirlo. Volvamos a pasar la cinta miles de veces y hallaremos las buenas mañas una y otra vez, en un linaje o en otro. La evolución convergente no es evidencia de progreso global, pero es una evidencia apabullante del poder de los procesos de la selección natural. Este es el poder de los subyacentes algoritmos, desarrollándose a lo largo de las vías evolutivas en sentido descendente, aunque, gracias a las grúas que han construido a lo largo de estos caminos, admirablemente capaces de descubrimiento, reconocimiento y sabia decisión. No hay espacio para los ganchos celestes, ni necesidad de ellos.

¿Acaso Gould pensó que su tesis de la contingencia radical refutaría la idea nuclear de Darwin, es decir, la evolución como un proceso algorítmico? Esta es, en principio, mi conclusión. Los algoritmos, en la imaginación popular, son algoritmos *para* producir un resultado concreto. Como ya he dicho en el capítulo 2, la evolución puede ser un algoritmo, una evolución que nos ha producido por un proceso algorítmico, sin que sea verdad que la evolución es un algoritmo *para* producirnos. Pero si alguien no comprende este punto, podría pensar: «Si no somos el resultado previsible de la evolución, la evolución no puede ser un proceso algorítmico».

Y quien esto piensa estaría muy motivado para probar la «contingencia radical», si desea demostrar que la evolución no ha sido un proceso algorítmico. En este asunto puede no haber ganchos celestes reconocibles, pero, al menos, sabremos si todo no tuvo lugar nada más que con grúas.

¿Es posible que Gould estuviera tan confundido acerca de la naturaleza de los algoritmos? Como veremos en el capítulo 15, Roger Penrose, uno de los más distinguidos matemáticos mundiales, escribió un extenso libro sobre las máquinas de Turing, los algoritmos y la imposibilidad de la inteligencia artificial; y todo el libro está basado en esa confusión. Esto no es realmente, por parte del pensador, un error inverosímil. Una persona a la que realmente no le gusta la peligrosa idea de Darwin con frecuencia tiene dificultades para colocar la idea en el enfoque apropiado.

Con esto concluye mi «historia» (como aquellas «al fin y al cabo historias» de Kipling) acerca de cómo Stephen Jay Gould se convirtió en el niño que gritó que viene el lobo. Sin embargo, un buen adaptacionista no debe quedar satisfecho con una historia verosímil. Al menos, debe hacer un esfuerzo para considerar, y descartar, hipótesis alternativas. Como he dicho al comienzo, estoy más interesado en las razones que han mantenido el mito que en los motivos reales del hombre real, pero puede parecer poco sincero por mi parte no mencionar las obvias explicaciones «rivales» que piden ser consideradas: política y religión. (Es posible que existan motivaciones políticas o religiosas detrás del ardiente deseo de ganchos celestes que yo imputo a Gould, pero éstas no serían hipótesis rivales; serían elaboraciones de mi interpretación, que podrían posponerse para otra ocasión. Aquí debo considerar brevemente si una de éstas —política o religión— puede ofrecer una más simple y directa interpretación de su campaña antidarwinista, obviando mi análisis. Muchos de los críticos de Gould piensan que sí; yo creo que están pasando por alto la posibilidad más interesante).

Gould nunca ha hecho un secreto de sus ideas políticas. Aprendió el marxismo de su padre, según nos dice, y hasta hace muy poco se ha dejado oír bastante y ha sido muy activo en la política de la extrema izquierda. Muchas de sus campañas contra científicos concretos y específicas escuelas de pensamiento dentro de la ciencia han sido llevadas en términos explícitamente políticos —en realidad, explícitamente marxistas— y a menudo han tenido como objetivos a pensadores de la extrema derecha. Así, no es de extrañar que sus oponentes y críticos hayan supuesto a menudo, por ejemplo, que su puntuaciomsmo era expresión de su antipatía marxista por la reforma que tenía lugar en la biología. Los reformadores son los peores enemigos de los revolucionarios, como todos sabemos. Pero esto, en mi opinión, es tan sólo una lectura superficialmente plausible de las razones de Gould. Después de todo, John Maynard Smith, su polo opuesto en las controversias sobre la evolución, tiene un trasfondo marxista tan rico y activo como el de Gould, y hay otros que simpatizan con la extrema izquierda contra quienes Gould ha dirigido ataques. (Y además están todos los liberales de la ACLU, como yo mismo, aunque dudo si él los conoce o les importa). Al regresar de una visita a Rusia Gould (1992b) dirigió su atención, como había hecho antes a menudo, a la diferencia entre la gradualidad de la reforma y la instantaneidad de la revolución. En esta interesante obra, Gould refleja sus experiencias en Rusia y el fracaso del marxismo allí —«Sí, la realidad rusa desacredita una economía específicamente marxista»— pero sigue diciendo que, no obstante, Marx

estaba en lo cierto en lo tocante a «la validez del modelo a mayor escala del cambio puntuado» (1992b:14). Esto no significa que para Gould la teoría económica y social de Marx no fuera nunca la cuestión, aunque no es difícil creer que Gould mantenía a bordo sus actitudes acerca de la evolución, mientras arrojaba el lastre de algún equipaje político que había sobrevivido a su bienvenida.

En cuanto a la religión, mi propia interpretación es, en un importante sentido, una hipótesis sobre los anhelos religiosos de Gould. Su antipatía hacia la peligrosa idea de Darwin la interpreto fundamentalmente como un deseo de proteger o restablecer la idea de la primera mente, de la visión desde arriba hacia abajo de John Locke, al menos para asegurar *nuestro* lugar en el cosmos con un gancho celeste. (El humanismo secular es una religión para algunos, quienes a veces creen que la humanidad no puede ser lo bastante especial para preocuparse de que pueda ser un producto de un proceso meramente algorítmico, un tema que analizaré en los últimos capítulos). Gould ha considerado, ciertamente, que su tarea tiene implicaciones cósmicas, algo que aparece especialmente claro en las epifanías acerca de Burgess Shale en su libro *La vida maravillosa*. Esto convierte, en un importante sentido, su visión del mundo en una cuestión religiosa, tenga o no entre sus directos ancestros el credo oficial de su herencia religiosa, o cualquier otra religión organizada. Gould cita con frecuencia la Biblia en su columna mensual y, a veces, su efecto retórico es impresionante. De un artículo con esta sentencia inicial, lo más seguro es que uno piense que ha sido escrito por un hombre religioso: «Del mismo modo que el Señor tomó la totalidad del mundo en sus manos, nosotros ansiamos incluir toda una cuestión en un epigrama inteligente» (Gould 1993e:4).

A menudo Gould ha afirmado que no hay conflicto entre la teoría de la evolución y la religión.

> A menos que la mitad de mis colegas sean estúpidos —no puede haber conflicto— sobre bases primarias y empíricas —entre ciencia y religión. Conozco cientos de científicos que comparten una convicción acerca del hecho de la evolución y la enseñan de la misma forma. Entre éstos encuentro un completo espectro de actitudes religiosas, desde la devota plegaria y la adoración diaria hasta el más absoluto ateísmo. Por lo tanto, no hay correlación entre la creencia religiosa y la confianza en la evolución, a menos que la mitad de esta gente esté loca (Gould 1987:68).

Ciertas alternativas más realistas serían que aquellos evolucionistas que no ven conflicto entre la evolución y sus creencias religiosas han tenido el cuidado de no observar el problema tan de cerca como nosotros lo hemos estado observando, o bien la de los que mantienen unos puntos de vista

religiosos que conceden a Dios el desempeño de lo que pudiera denominarse un papel meramente ceremonial (en el capítulo 18 hay más discusión sobre este punto). O quizá, como Gould, se cuidan de delimitar el presumible papel de ambas, ciencia y religión. La compatibilidad que Gould observa entre ciencia y religión se mantiene sólo mientras la ciencia conoce su lugar y declina ocuparse de las grandes cuestiones. «La ciencia no se ocupa de las interrogantes en torno a los más remotos orígenes» (Gould 1991b:459). Un modo de interpretar las campañas de Gould dentro de la biología a lo largo de los años puede ser como un intento de restringir la teoría de la evolución a una tarea modesta, creando un *cordón sanitario* entre ella y la religión. Gould dice, por ejemplo:

La evolución, de hecho, no es el estudio de los orígenes. Incluso la más restringida (y científicamente permisible) cuestión acerca del origen de la vida en la Tierra queda fuera de este dominio. (Sospecho que este interesante problema cae primariamente dentro de la competencia de la química y la física de los sistemas autoorganizados). La evolución estudia las vías y mecanismos de los cambios orgánicos siguiendo al origen de la vida (Gould 1991b:455).

Este punto de vista dejaría todo el tema del capítulo 7 fuera de los límites de la teoría de la evolución pero, como hemos visto, ésta ha llegado a ser el verdadero fundamento de la teoría de Darwin. Gould parece creer que debe desalentar a sus colegas evolucionistas de intentar extraer grandes conclusiones filosóficas de sus trabajos pero, de este modo, está tratando de negar a los otros lo que se permite a sí mismo. En la frase que cierra su libro *La vida maravillosa* Gould está dispuesto a presentar una conclusión específicamente religiosa de sus propias consideraciones sobre la implicación de la paleontología:

Somos los vástagos de la historia, y debemos establecer nuestros propios caminos en el más diverso e interesante de los universos concebibles: un universo indiferente a nuestros sufrimientos y que, en consecuencia, nos ofrece la máxima libertad para prosperar o fracasar en la vía que hayamos escogido (Gould 1989a:323).

Es bastante curioso que estas palabras me impresionen como una hermosa expresión de las implicaciones de la peligrosa idea de Darwin, en absoluto en conflicto con la idea de que la evolución es un proceso algorítmico. Esta es, ciertamente, una opinión que comparto de todo corazón. Sin embargo, Gould parece creer que la visión de mundo que está combatiendo tan vigorosamente es determinista y a histórica, en conflicto con este credo de libertad. El «hiperdarwinismo», el fantasma de Gould, es simplemente la tesis según la cual no son necesarios, en ningún momento, ganchos celestes para explicar la

tendencia que tienen las ramas del árbol de la vida a elevarse. Como otros antes que él, Gould ha tratado de demostrar la existencia de saltos, aceleraciones y otras inexplicables trayectorias, inexplicables para las herramientas del «hiperdarwinismo». Pero, a pesar de lo «radicalmente contingentes» que estas trayectorias puedan haber sido, a pesar de lo «puntuado» que el camino de la evolución haya sido, sea con saltos «no darwinianos» o con insondables «mecanismos de especiación», nada de esto crea más libertad de movimientos para que «el poder de los acontecimientos contemporáneos y las características personales conformen y dirijan el camino real seguido entre miles y miles de posibilidades». No era necesaria más libertad de movimientos (Dennett 1984).

Un efecto impresionante de la campaña de Gould a propósito de la contingencia fue que terminó poniendo a Nietzsche patas arriba. Cabe recordar que Nietzsche pensó que nada podía ser más terrorífico, más destructivo para el mundo que pensar que si pudiéramos volver a pasar la cinta, todo sucedería una y otra vez: el eterno retorno, la idea más enfermiza que alguien pudo haber concebido nunca. Nietzsche entendió como misión suya enseñar a la gente a decir «¡sí!» a esta terrible verdad. Gould, por su parte, pensó que debía mitigar el terror de la gente con la negación de esta idea: aunque la cinta de la vida siguiera pasando una y otra vez, *¡nunca podría suceder esto de nuevo!* ¿Son ambas proposiciones igualmente disparatadas para la mente[111]? ¿Cuál es peor? ¿Que sucediera una y otra vez o bien que nunca sucediera de nuevo? Bien, podría decir Tinker, si sucediera o no otra vez, no se estaría negando; de hecho la verdad es un mezcla de ambas: un poquito de suerte, un poquito de alguna vez. Esta es la peligrosa idea de Darwin, guste o no guste.

CAPÍTULO 11
Controversias implícitas

I. *Una nidada de herejías inofensivas*

> Encuentro, al releerlo, que el cuadro que presenta es muy semejante al que yo pintaría si tuviera que comenzar ahora y escribir un libro totalmente nuevo.
>
> John Maynard Smith, *The Theory of Evolution*

Antes de volver a examinar, en la parte III, la peligrosa idea de Darwin aplicada a la humanidad —y a las humanidades— permítaseme hacer una pausa para inventariar el resultado de nuestra revisión de las controversias surgidas dentro de la propia biología. Gould ha hablado del «endurecimiento» de la síntesis moderna del darwinismo, pero también ha mostrado su frustración a causa de que la síntesis moderna se mantiene en continuo movimiento ante sus propios ojos, lo que hace tan difícil que acierte con un buen disparo. Los defensores de la síntesis moderna siguen cambiando la historia, dando opciones a los revolucionarios al incorporar en la propia síntesis aquellos puntos de vista que les parecen buenos. ¿Hasta qué punto es segura la síntesis moderna —o su sucesora innominada— si pensamos que ha cambiado demasiado como para mantener su viejo título? ¿Es el cuerpo de doctrina del darwinismo actual demasiado duro o demasiado blando? Como sucedió con la cama favorita de Ricitos de Oro (la niña rubia de ojos azules escogió —en el famoso cuento infantil— la del osezno, desechando la del oso padre y la de la osa madre), su elección ha demostrado ser la correcta: el darwinismo ha sido duro cuando tenía que serlo, y condescendiente en aquellos temas que permanecen abiertos a ulteriores investigaciones y debates.

Para tener una idea apropiada de lo que es duro y de lo que es blando, debemos retroceder un poco y revisar todo el panorama. A algunos les gustaría destruir las credenciales de la peligrosa idea de Darwin, y nosotros les podemos ayudar puntualizando las controversias en las cuales no necesitan gastar sus energías, ya que, no importa lo que resulte, la idea de Darwin

sobrevivirá intacta o fortalecida. Y después también podemos señalar aquellos puntos firmes y fijos, los cuales, de ser destruidos, realmente se llevarían por delante al darwinismo; no obstante, dichos puntos permanecen firmes y fijos por buenas razones, y son tan difíciles de mover como las pirámides.

Permítaseme considerar, en primer lugar, algunos intentos de herejías que no arrasarán al darwinismo, ni siquiera si fueran confirmadas. Probablemente la más conocida ha sido la defendida, en años recientes, por el astrónomo rebelde Fred Hoyle, quien argumenta que la vida no se originó —podía no haberse originado— en la Tierra, sino que fue «sembrada» desde el espacio exterior (Hoyle 1964, Hoyle y Wickramasinghe 1981). Francis Crick y Leslie Orgel [1973] y Crick [1981] puntualizaron que la idea de la *panspermia* había sido propugnada en varias formas desde el principio de siglo, cuando Arrhenius [1908] acuñó el término, y a pesar de lo improbable que pueda ser, no es una idea incoherente. No se ha refutado todavía que las primitivas formas de vida (algo tan «simple» como un macro o tan complejo como una bacteria) llegaran en un asteroide o en un cometa, procedentes de alguna otra región del universo para colonizar nuestro planeta. Crick y Orgel dieron un paso más: es posible que la *panspermia* fuera dirigida, que la vida comenzara en la Tierra como resultado de que nuestro planeta hubiera sido *deliberadamente* «infectado» o colonizado por formas de vida procedentes de algún otro lugar del universo con una ventaja de salida sobre nosotros y, por tanto de modo indirecto, somos sus productos. Dado que ahora podemos enviar una nave espacial cargada de formas de vida a otro planeta —y podemos, pero no debemos hacerlo— entonces, por paridad del razonamiento, otros podrían haberlo hecho. Desde que Hoyle [1964:43] —al contrario que Crick y Orgel— lanzó la sospecha de que, a menos que la panspermia sea verdad, «la vida tiene poco significado, y debe ser juzgada como un mero capricho cósmico», no es sorprendente que muchos, incluyendo al propio Hoyle, hayan supuesto que la panspermia, de ser confirmada, destrozaría al darwinismo, esta temida amenaza para el significado de la vida. Y dado que la panspermia es a menudo tratada irrisoriamente por los biólogos —a Hoyle le han llamado el Aullador— se ha fomentado la ilusión de que en ella subyace una grave amenaza que golpea al propio núcleo del darwinismo.

Nada puede estar más lejos de la verdad. El propio Darwin conjeturó que la vida comenzó sobre la Tierra en una charca pequeña y caliente, pero que igualmente podía haber comenzado en alguna caliente y sulfurosa olla a presión situada bajo tierra (como ha sido recientemente propuesto por Stetter

y otros 1993) o, en cuanto a esta idea se refiere, en algún otro planeta desde donde viajó hasta aquí después de que alguna colisión astronómica pulverizara su lugar de nacimiento. Donde quiera y cuando quiera que la vida comenzara, tendría que hacerlo por sí misma, sin ayuda ajena, mediante alguna versión del proceso que hemos analizado en el capítulo 7; éste es el punto en el que insiste el darwinismo ortodoxo. Y, como ha señalado Manfred Eigen, la panspermia no resuelve el difícil problema de cómo tuvo lugar esta puesta en marcha sin ayuda ajena: «La discrepancia entre los números de secuencias que pueden ser sometidas a prueba en la práctica, e imaginables en teoría, es tan grande que los intentos de explicación a través del desvío del origen de la vida a otros lugares, de la Tierra al espacio exterior, no ofrecen una solución aceptable para el dilema. La masa del universo es "sólo" 10^{29} veces y su volumen "sólo" 10^{57} veces mayor que la masa y el volumen de la Tierra» (Eigen 1992:11).

La razón de que la ortodoxia prefiera asumir un lugar de nacimiento sobre la Tierra es que ésta es la hipótesis más simple y más accesible científicamente. Pero esto no lo convierte en verdadera. Cualquiera cosa que sucediera, sucedió. Si Hoyle está en lo cierto, entonces —¡maldita sea!— será mucho más difícil confirmar o rechazar cualquier hipótesis detallada acerca de cómo comenzó exactamente la vida. La hipótesis de que la vida comenzó en la Tierra tiene la virtud de disponer de algunas limitaciones admirablemente restrictivas para los que cuentan cuentos: la globalidad del cuento ha de desarrollarse durante cinco mil millones de años y ha de comenzar con condiciones conocidas como las que existieron en la Tierra en los primeros días. A los biólogos *les gusta* verse obligados a trabajar dentro de estas limitaciones; *desean* plazos temporales y una corta lista de materias primas, mientras más exigente mejor**[112]**. De este modo esperan que nunca será confirmada una hipótesis que abra posibilidades tan vastas que sean casi imposibles de evaluar en detalle. Los argumentos que Hoyle y otros han presentado a favor de la panspermia pertenecen todos al *phylum* de los que repiten eso de «por otro lado no hay bastante tiempo» y los teóricos de la evolución prefieren mantener intactos los plazos temporales y buscar más grúas que permitan hacer toda la elevación evolutiva en el tiempo disponible. De hecho, esta política ha proporcionado resultados excelentes. Si la hipótesis de Hoyle llegara a confirmarse algún día, sería un día triste para los teóricos de la evolución, no porque el darwinismo quedara destruido, sino porque haría que ciertas características importantes del darwinismo fueran menos confirmables, más especulativas.

Por la misma razón, los biólogos serían hostiles ante cualquier hipótesis que propusiese que el antiguo ADN ha sido manipulado por acopladores de genes de otro planeta que hubieran llegado a poseer, antes que nosotros, una tecnología muy avanzada y nos la hubieran jugado. Los biólogos se mostrarían hostiles a esta hipótesis, pero pasaríamos tiempos difíciles para desaprobarla. Todo esto plantea una importante cuestión acerca de la naturaleza de la evidencia en la teoría de la evolución que merece la pena ser analizada con mayor detalle, con la ayuda de algunos experimentos mentales (Dennett 1987b, 1990b).

Como han señalado muchos comentaristas, las explicaciones evolucionistas son, ineludiblemente, narraciones históricas. Ernst Mayr lo expresó de esta manera: «Cuando uno intenta explicar las características de algo que es el producto de la evolución, debe intentar reconstruir la historia evolutiva de estas características» (1983:325). Pero los hechos históricos concretos tienen un papel esquivo en tales explicaciones. La teoría de la selección natural muestra que todas las características del mundo natural *pueden* ser, en último término, el producto de un proceso de reproducción diferencial durante largos períodos de tiempo, que sería mecánico, ciego, imprevisible y, desde luego, no teleológico. Aunque algunas características del mundo natural —las cortas extremidades del perro salchicha y el ganado vacuno para carne de la raza Black Angus, la gruesa piel de los tomates— son productos de la selección artificial, realmente el objetivo del proceso y las razones del diseño desempeñan un papel en su desarrollo. En estos ejemplos, el objetivo estaba explícitamente representado en las mentes de los criadores que hicieron la selección. Así, la teoría de la evolución pudo hacer posible la existencia de tales productos y de tales procesos históricos, como casos especiales; son organismos diseñados con la ayuda de supergrúas. Surge ahora la pregunta: ¿pueden distinguirse estos casos en los análisis retrospectivos?

Imaginemos un mundo en el cual manos *reales* de otra galaxia complementarán a las «manos ocultas» de la selección natural. Imaginemos que la selección natural en este planeta hubiera sido ayudada e incitada, a través de los tiempos, por visitantes: diseñadores de organismos chapuceros, sagaces y representantes de la razón, como los criadores y cultivadores de animales y plantas de nuestro mundo real, pero no limitados a organismos «domesticados», diseñados para uso humano. (Para hacer esta imagen más gráfica, podemos suponer que han manipulado la Tierra como un «parque temático», creando *phyla* completos con propósitos educativos y de

entretenimiento). Estos bioingenieros realmente habrían formulado, representado y actuado desde el punto de vista de la racionalidad de sus diseños, del mismo modo que los ingenieros automovilísticos o los actuales manipuladores de genes. Permítasenos suponer, entonces, que están escondidos. ¿Sería detectable su trabajo para los biólogos actuales mediante algún tipo de análisis?

Si nos encontramos con que algunos organismos llevan consigo su manual de uso, esta sería una vía muerta, ya que la mayor parre del ADN de cualquier genoma permanece sin expresar y habitualmente se denomina «ADN chatarra»; NovaGene, una compañía de biotecnología afincada en Houston, ha encontrado un uso para este ADN y han adoptado la política del «ADN marcado»; escribiendo los codones más cercanos de la marca de su compañía en el ADN chatarra de los aminoácidos, asparagina, glutamina, valina, alanina, glicina, ácido glutámico, asparagina, ácido glutámico = NQVAGENE (publicado en *Scientific American*, junio de 1986:70-71). Todo esto sugiere un nuevo ejercicio de «traducción radical» (Quine 1960) para los filósofos: ¿cómo, en principio o en la práctica, podemos confirmar o no la hipótesis de que las marcas registradas —o los manuales de uso u otro tipo de mensajes— fueran discernióles en el ADN chatarra de cualquier especie? Hace tiempo que la presencia de ADN no funcionante en el genoma no se considera como un rompecabezas. La teoría del gen egoísta de Dawkins [1976] predice este hecho y la elaboración de la idea del ADN egoísta fue simultáneamente desarrollada por Doolittle y Sapienza [1980] y Orgel y Crick [1980] (véase Dawkins 1982:cap. 9, para más detalles). Esto no demuestra, sin embargo, que el ADN chatarra *no pueda tener* una función más importante y, en consecuencia, poseer, después de todo, un significado. Nuestros imaginados intrusos intergaláticos podrían reciclar tan eficientemente el ADN chatarra para sus propios propósitos como la compañía NovaGene lo recicla para los suyos.

Encontrar la marca de la aplicación de esa alta tecnología, como se encuentra el legendario *graffiti* norteamericano que recuerda que «Kilroy estuvo aquí», escrita en el genoma de una col o de un rey sería desconcertante, pero ¿qué pasaría si no hubieran dejado esparcidas tales claves? ¿Podría un examen más minucioso del diseño de los propios organismos —los fenotipos— revelar algún indicio de discontinuidades en el proceso? El acoplamiento genético [ingeniería genética] es la grúa más poderosa que se ha descubierto. ¿Hay diseños que sencillamente no pueden elevarse sin la ayuda de una grúa en concreto? Si hay diseños que no se

pueden conseguir a través de un proceso gradual, paso a paso, en el cual cada paso es, al menos, no peor para las oportunidades de supervivencia de los genes que el anterior, entonces la existencia de tal diseño en la naturaleza podría requerir, en algún punto de su línea de ancestros, la ayuda de la mano de un diseñador previsor, sea un acoplador de genes o un criador, quien, de algún modo, mantendría la necesaria sucesión de descarriados intermediarios hasta obtener la buscada progenie. Pero ¿podemos establecer de modo concluyente que algún diseño tendría esta característica si se hubiera *requerido* tal salto en sus ancestros? Durante un siglo, los escépticos han estado a la caza de tales casos, pensando que, si encontraban alguno, habrían refutado el darwinismo de manera concluyente: hasta ahora sus esfuerzos han mostrado una sistemática debilidad.

Consideremos el ala, que es el ejemplo más familiar. Las alas no pueden evolucionar de un solo golpe, dice el argumento escéptico estándar; y si nosotros, como darvinistas, imaginamos que las alas evolucionan gradualmente, debemos admitir que las alas incompletas no sólo no aportarían ningún valor, sino que habrían sido un estorbo. Los darwinistas no necesitamos admitir tal cosa. Las alas que sirven solamente para planear (pero no para un vuelo poderoso) han puesto de manifiesto netos beneficios para muchas criaturas reales, e incluso siendo más cortas, estas protuberancias menos efectivas aerodinámicamente pudieron haber evolucionado, por alguna u otra razón, y han sido recicladas. Muchas versiones de este relato —y otros cuentos— se han contado para llenar este hueco. Las alas no son una dificultad para los darvinistas ortodoxos, y si lo son, son dificultades de ricos. ¡Hay muchos y diferentes modos verosímiles de contar la historia de cómo las alas funcionantes pudieron haber evolucionado gracias a un incremento gradual! Esto demuestra lo difícil que sería para cualquiera el diseño de un argumento insuperable para probar que una característica concreta debe haber surgido por un salto, pero, al mismo tiempo, demuestra que sería igualmente difícil probar que una característica determinada pudo haber aparecido sin un salto, *sin ayuda* de manos humanas u otro tipo de manos inteligentes.

Todos los biólogos a los que les he preguntado sobre este punto han coincidido conmigo en que no existen marcas seguras de la selección natural, en oposición a lo que sucede en la artificial. En el capítulo j nos ocupamos del concepto de posibilidad e imposibilidad biológica estricta para una noción graduada de la probabilidad biológica, pero, incluso en estos términos, no está claro cómo se pueden valorar los organismos como «probables» o «muy probables» o «extremadamente probables» productos de la selección artificial.

¿Debe considerarse esta conclusión como una terrible perplejidad para los evolucionistas en su lucha contra los creacionistas? Podemos imaginar los titulares: «¡Los científicos aceptan que la teoría de Darwin no puede desaprobar el diseño inteligente!». Sin embargo, sería temerario, para cualquier defensor del neodarwinismo, sostener que la teoría contemporánea de la evolución nos da el poder de interpretar la historia de modo tan preciso a partir de los datos actuales, como descartar la presencia, más temprana históricamente, de diseñadores racionales, un fantasía totalmente inverosímil, aunque una posibilidad al fin y al cabo.

En el mundo donde vivimos, hay organismos que *sabemos* que son el resultado de esfuerzos de rediseño, previsores y dirigidos a un objetivo, pero este conocimiento depende, a su vez, de nuestro conocimiento directo de los acontecimientos históricos recientes; hemos observado realmente cómo trabajan los criadores. Estos acontecimientos especiales no darán lugar, probablemente, a ninguna sombra de fósiles en el futuro. Optando por una variante más simple de nuestro experimento mental, podemos suponer que enviamos a los biólogos «marcianos» una gallina ponedora de huevos, un perro pequinés, una golondrina de los graneros y un leopardo, y les rogamos a estos biólogos que averigüen qué diseños llevan la marca de una intervención de seleccionadores artificiales. ¿En qué pueden confiar los biólogos marcianos? ¿Cómo pueden argumentarlo? Pueden detectar que la gallina no cuida «de modo apropiado» sus huevos; algunas variedades de gallinas han perdido su instinto de empollar, por lo que se extinguirían muy pronto si no fuera por el ambiente creado por las incubadores artificiales, proporcionadas por los seres humanos.

Los biólogos pueden notar que el perro pequinés está patéticamente mal dotado para defenderse por sí mismo en cualquier ambiente exigente que pudieran imaginar. Pero la innata atracción de las golondrinas de los graneros por anidar en el interior de las construcciones, pudiera volverlos locos, y hacerles pensar que se trata de algún tipo de animal doméstico, y cualquier característica del *cheetah* les convencería de que es una criatura con el salvaje poder también presente en los galgos, características que sabemos que han sido pacientemente estimuladas por los criadores. Después de todo, los ambientes artificiales son parte de la naturaleza, así que es improbable que haya *algún* signo claro de selección artificial que se pueda leer en un organismo, en ausencia de información procedente de una persona informada sobre la historia real que ha creado el organismo.

Una manipulación chapucera realizada por visitantes intergalácticos en el ADN de las especies terrestres no puede descartarse, a menos que se trate de una fantasía completamente gratuita. Hasta ahora, no se ha encontrado nada en la Tierra que sugiera que esta hipótesis merece ser investigada. Y recordemos —me apresuro a añadir, no sea que los creacionistas se animen— que aunque fuéramos a descubrir y trasladar el «mensaje de marca registrada» en nuestro escaso ADN, o encontrar alguna otra marca incontestable de tal alteración precoz, esto no haría nada para rescindir la pretensión de la teoría de la selección natural de explicar todo el diseño en la naturaleza sin la invocación de un previsor Creador-Diseñador situado *fuera del sistema*. Si la teoría de la evolución mediante selección natural puede explicar la existencia de gente en la compañía NovaGene que inventa un ADN con marca de fábrica, también puede dar cuenta de la existencia de predecesores cuyas firmas están a nuestro alrededor para ser descubiertas.

Conocida esta posibilidad, no obstante lo improbable que parezca, también observamos que si los escépticos no han encontrado nunca su Santo Grial, condensado en la expresión «No podrías llegar aquí desde allí», a lo largo de una presunta línea evolutiva, se trate de un órgano o de un organismo, después de rodo, este hecho no hubiera sido *concluyente* contra el darwinismo. ¡El mismo Darwin dijo que tendría que abandonar su teoría si este fenómeno se descubriera (véase la nota 5 del capítulo 2), pero ahora podemos comprobar que siempre hubiera sido lógicamente coherente para los darwinianos (no obstante ser inaceptable y *ad hoc*) replicar que lo que estaban mostrando era evidencia para la sorprendente hipótesis de los intrusos intergalácticos! El poder de la teoría de la selección natural no es el poder de probar exactamente cómo era la (pre)historia, sino tan sólo el poder de probar cómo podría haber sido. Teniendo en cuenta lo que sabemos acerca de cómo son las cosas.

Antes de abandonar este curioso tema de las nada bienvenidas, aunque, en todo caso, no fatales herejías, permítaseme considerar una que es un poco más realista. ¿Surgió la vida en la Tierra solamente una vez o muchas veces? La ortodoxia supone que sucedió una vez, pero no pasa nada si, en efecto, la vida hubiera surgido dos veces o diez o cientos de veces. No obstante lo improbable que el forzado acontecimiento inicial pudiera haber sido, no debemos compartir la falacia del jugador que supone que después de que algo sucedió una vez, aumentan las probabilidades contrarias a que suceda de nuevo. Aún más, la cuestión de cuántas veces surgió la vida independientemente abre algunas interesantes perspectivas. Si al menos

algunas de las disposiciones espaciales en el ADN son puramente arbitrarias, entonces ¿no puede haberse tratado de dos lenguajes genéticos diferentes coexistiendo, como el francés y el inglés, sin relación alguna entre ellos? Esta posibilidad no se ha descubierto: el ADN ha coevolucionado claramente con su precursor el ARN, pero esto no demuestra que la vida no surgiera más de una vez, ya que todavía no conocemos cómo era realmente la amplitud del espectro de variaciones en el código genético.

Supongamos que hubiera exactamente dos lenguajes igualmente viables y construibles del ADN, el «Mendelese» (el nuestro) y el «Zendelese». Si la vida surgió dos veces, tendríamos cuatro posibilidades equiprobables: dos veces el lenguaje «Mendelese», dos veces el lenguaje «Zendelese», «Mendelese» y «Zendelese» o «Zendelese» y «Mendelese». Si pasáramos la cinta de la vida muchas veces y observáramos las ocasiones en las que la vida surge dos veces, cabría esperar que la mitad del tiempo aparecieran ambos lenguajes, pero en un cuarto de aquellos *replay* sólo el «Mendelese» aparecería. En aquellos mundos, el lenguaje del ADN de todos los organismos sería el mismo, aunque otro lenguaje fuera posible. Esto demuestra que la «universalidad» (al menos en nuestro planeta) del lenguaje del ADN no permite la válida inferencia de que todos los organismos han surgido de un único progenitor, el último Adán, dado que, *ex hipothesi* en estos casos, Adán pudo haber tenido un gemelo enteramente independiente, compartiendo accidentalmente el mismo lenguaje del ADN. Naturalmente, si la vida surgió muchas más veces —es decir, cientos de veces— bajo las mismas condiciones, entonces la probabilidad de que sólo uno de los dos equiprobables lenguajes aparecieran desciende en vertical, hasta desvanecerse. Y si de hecho hay más de dos códigos genéticos, igualmente utilizables, esto cambiaría las implicaciones acerca de la probabilidad. Pero hasta que conozcamos mejor la amplitud de las posibilidades genuinas y sus probabilidades asociadas, no dispondremos de ninguna buena palanca que nos permita decidir, con seguridad, que la vida surgió una vez. Por el momento, esta es la hipótesis más simple: que la vida sólo ha surgido una vez.

2. *Tres perdedores: Teilhard, Lamarck y la mutación dirigida*

Ahora, vayamos al extremo opuesto y consideremos una herejía que sería verdaderamente fatal para el darwinismo si no se tratase de una alternativa tan confusa y, en último término, contradictoria en sí misma: el intento del jesuita paleontólogo Teilhard de Chardin de reconciliar su religión con su creencia en

la evolución. Teilhard propuso una versión de la evolución que ponía a la humanidad en el centro del universo y describió al cristianismo como una expresión del objetivo, el «punto Omega», hacia el cual se dirige toda la evolución. Teilhard consiguió un espacio para el pecado original (en su versión ortodoxa católica, no en la versión científica que presenté en el capítulo 8). Para su decepción, la Iglesia consideró su propuesta como una herejía y le prohibió enseñar en París, por lo que pasó el resto de sus días en China, estudiando fósiles, hasta su muerte en 1955. Su libro *El fenómeno humano* (1959), publicado tras su muerte, obtuvo una acogida internacional, aunque el *establishment* científico, el darwinismo ortodoxo en particular, fue tan categórico como la Iglesia rechazándolo como herético. Es de justicia decir que en los años que siguieron a la publicación de su obra se ha hecho evidente, hasta el punto de alcanzar la unanimidad entre los científicos, que Teilhard no ofreció nada serio como alternativa a la ortodoxia; sus ideas eran confusas y el resto era una redescripción rimbombante de la ortodoxia[113]. El clásico intento de salvación corrió a cargo de sir Peter Medawar y se encuentra en su libro de ensayos *Pluto's Republic*. Como muestra una frase: «A pesar de todos los obstáculos que Teilhard, quizá sabiamente, había puesto en nuestro camino, es posible discernir una secuencia de pensamientos en *El fenómeno humano*» (1982:245).

El problema con la idea de Teilhard es simple. Niega enfáticamente la idea fundamental: evolución es un proceso algorítmico, no inteligente y sin propósito. El suyo no era un compromiso constructivo, sino una traición al punto de vista central que permitió a Darwin desplazar la idea de la mente primera de Locke. Alfred Russel Wallace tuvo la tentación del mismo abandono de la idea central, como hemos visto en el capítulo 3, pero Teilhard cae en ella con todo su corazón y lo convierte en la pieza central de su visión alternativa[114]. La estima en que los no científicos tienen todavía al libro de Teilhard, el tono respetuoso con el que son aludidas sus ideas, es testimonio de cuán profundamente se aborrece la peligrosa idea de Darwin, un aborrecimiento tan grande que justifica cualquier ausencia de lógica y tolera cualquier opacidad en lo que se presenta como una argumentación, siempre que su línea fundamental prometa aliviar las presiones del darwinismo.

¿Y qué decir de otra notoria herejía, el lamarckismo, es decir, la creencia en el carácter hereditario de las características adquiridas[115]? Aquí la situación es bastante más interesante. La principal atracción del lamarckismo ha sido siempre su promesa de acelerar la travesía de los organismos por el espacio de diseño sacando ventaja de las mejoras adquiridas en el diseño por

organismos individuales durante sus vidas. ¡Tanto trabajo de diseño por hacer y tan poco tiempo! Pero la perspectiva del lamarckismo como *alternativa* al darwinismo puede descartarse simplemente desde el punto de vista lógico: la capacidad para poner en marcha la herencia lamarckiana *presupone*, en primer lugar, un proceso darwiniano (o un milagro) (Dawkins 1986a:299-300). Pero ¿no podría la herencia lamarckiana ser una importante grúa *dentro* de la estructura darwiniana? El propio Darwin, como es notorio, incluyó en su versión a la herencia lamarckiana de la evolución como un proceso amplificador, sumado a la selección natural. Darwin estaba en condiciones de considerar la idea lamarckiana, dados sus nebulosos conocimientos de los mecanismos de la herencia. (Para hacernos una idea de lo poco constreñida que era la imaginación de Darwin con respecto a los mecanismos de la herencia, consúltese a Desmond y Moore 1991:53 ISS., con una revisión de la atrevida especulación de Darwin acerca de la «pangénesis»).

Una de las contribuciones fundamentales al neodarwinismo, después del propio Darwin, fue la firme distinción de August Weismann [1893] entre la *línea germinal* y la *línea somática*; la línea germinal está constituida por las células sexuales en los ovarios o gónadas del organismo mientras que las otras células del cuerpo pertenecen a la línea somática. Lo que le sucede a la línea de células somáticas durante su vida tiene una relevancia, naturalmente, como lo tiene el que la línea germinal termine por fluir hacia una progenie, pero los cambios en las células somáticas mueren con aquellas células, mientras que solamente los cambios en las células germinales —es decir, las mutaciones— pueden seguir adelante. Esta doctrina, a veces llamada weissmanismo, es el baluarte que la ortodoxia levantó contra el lamarckismo, doctrina que el propio Darwin pensó que se debía contener. ¿Puede el weissmanismo aún ser destruido? Hoy las probabilidades contra el lamarckismo como una gran grúa en el proceso evolutivo parecen más formidables (Dawkins 1986a:288-303). Para que funcione el lamarckismo, la información acerca de las características adquiridas en cuestión habrá de trasladarse, de algún modo, desde la parte corporal revisada, el soma, hasta el huevo o el esperma, es decir, la línea germinal. En general, este envío de mensajes es totalmente imposible —no se han encontrado canales de comunicación que puedan conducir este tráfico— pero dejemos a un lado esta dificultad. El problema más profundo se relaciona con la naturaleza de la información en el ADN. Como hemos visto, nuestro sistema de desarrollo embrionario toma las secuencias del ADN como un receta, no como un plano. No hay un mapa, punto por punto, que establezca la correspondencia entre las

partes del cuerpo y las partes del ADN. Esto es lo que hace extremadamente improbable —en algunos casos imposible— que cualquier cambio adquirido en una parte del cuerpo (en un músculo o en un pico o, en el caso de la conducta, un circuito de control nervioso de cualquier tipo) se corresponda con un cambio discreto en el ADN del organismo. De este modo, aunque existiese una vía para enviar una orden de cambio a las células sexuales, no existiría una vía para *componer* la necesaria orden de cambio.

Véamos un ejemplo. La violinista desarrolla con asiduidad un magnífico *vibrato,* gracias fundamentalmente a los ajustes conseguidos en los tendones y ligamentos de su muñeca izquierda completamente diferentes de los que posee en su muñeca derecha, la de la mano que coge el arco. La receta para hacer una muñeca en el ADN humano produce ambas muñecas a partir de un solo conjunto de instrucciones que se aprovecha de la reflexión especular (a ello se debe que las muñecas sean tan parecidas) así que no habría un modo simple de cambiar la receta de la muñeca izquierda sin hacer el mismo cambio (no deseado) en la muñeca derecha. No es difícil imaginar cómo «en principio» el proceso embriológico podría estar atraído por la posibilidad de construir cada muñeca separadamente, después de la construcción inicial, pero incluso si este problema se pudiera solventar, serían mínimas las oportunidades de que fuera una mutación conseguida a través de la práctica, una revisión localizada y mínima en su ADN, que correspondiese casi totalmente a las mejoras que los años de ejercicio habían producido. De modo que, casi con toda certeza, los hijos de la violinista tendrán que aprender su *vibrato* del mismo modo que ella lo hizo.

Sin embargo, nada de esto es completamente concluyente, por lo que siguen brotando hipótesis en biología que tienen características, al menos fuertemente reminiscentes del lamarckismo y que a menudo se toman en serio, a pesar del tabú generalizado contra todo lo que huela a esta doctrina**[116]**. En el capítulo 3 llamé la atención sobre el hecho de que el efecto Baldwin ha sido a menudo pasado por alto e incluso desechado por los biólogos, que lo confunden con la temida herejía lamarckiana. La gracia salvadora del efecto Baldwin es que los organismos transmiten *su particular capacidad para adquirir* ciertas características, más que las características que realmente adquieren. Esto *tiene el efecto de* sacar ventaja durante las exploraciones de los diseños de los organismos individuales, como hemos visto, y, en consecuencia, se convierte en una poderosa grúa en las circunstancias apropiadas. Pero no es la grúa de Lamarck.

Finalmente, ¿cuál es la posibilidad de la mutación «dirigida»? Incluso desde Darwin, la ortodoxia ha supuesto que toda mutación es aleatoria; el ciego azar hace los candidatos. Mark Ridley nos proporcionó la declaración estándar:

> Se han propuesto varias teorías de la evolución mediante «variaciones dirigidas», pero debemos descartarlas. No hay evidencia de variación dirigida en la mutación, en la recombinación ni en el proceso de la herencia mendeliana. Cualquiera que sea la verosimilitud interna de estas teorías, son, de hecho, erróneas (1985:15).

Pero esto es un óbolo demasiado fuerte. La teoría ortodoxa *no debe presuponer* proceso alguno de mutación dirigida —lo que sería seguramente un gancho celeste— pero debe dejar abierta la posibilidad de que alguien descubriera mecanismos no milagrosos que puedan darle un sesgo a la distribución de mutaciones, que se acelera en determinadas direcciones. Las ideas de Eigen acerca de las cuasi-especies en el capítulo 8 son un ejemplo pertinente.

En capítulos anteriores he llamado la atención sobre otras posibles grúas que están siendo actualmente investigadas: el plagio, de especie a especie, de secuencias de nucleótidos (*Drosophila* de Houck), los entrecruzamientos hechos posibles por la innovación del sexo (algoritmos genéticos de Holland), la exploración de múltiples variaciones por pequeños equipos (*demes* de Wright) que retornan a la población progenitora («especies inteligentes» de Schull) y el «sorteo de las especies a alto nivel» de Gould, por nombrar cuatro. Dado que todos estos debates se acoplan con comodidad dentro de las amplias estancias del darwinismo contemporáneo, no necesitan un posterior escrutinio por parte de nosotros, por fascinantes que puedan ser. Casi siempre el tema en la teoría evolucionista no es, en principio, la posibilidad sino la relativa importancia, y los temas son siempre *bastante* más complejos de como yo los he presentado[117].

Sin embargo, hay un área de creciente controversia que merece un tratamiento más completo, no porque sea una amenaza para algo duro o quebradizo de la síntesis moderna del darwinismo —no obstante lo que resulte, el darwinismo seguirá siendo fuerte—, sino porque se ha observado que tiene implicaciones particularmente perturbadoras para la extensión en la humanidad del pensamiento evolucionista. Se trata del debate sobre las «unidades de selección».

3. «*Cui bono?*»

> Lo que es bueno para la General Motors es bueno para el país.
>
> Charles E. Wilson negó haberlo dicho, 1953

En 1952, Charles E. Wilson era presidente de la General Motors, y cuando fue reelegido Dwight Eisenhower como presidente de los Estados Unidos, lo nombró secretario de Defensa. En las sesiones de la Comisión de las Fuerzas Armadas del Senado, previas a su nombramiento, en enero de 1953, se le pidió a Wilson que vendiera sus acciones en la General Motors, pero puso objeciones. Cuando se le preguntó a Wilson si el mantenimiento de sus intereses económicos en la General Motors no podría influir indebidamente en sus juicios, replicó: «Durante años, pensé que lo que era bueno para el país era bueno para la General Motors, y viceversa». Desafortunadamente para él lo que realmente dijo no tuvo bastante poder de reproducción —aunque tuvo el suficiente para que yo pudiera localizar un descendiente en un libro de referencias y reproducirlo de nuevo en la frase anterior—. Por otro lado, lo que se reprodujo de su testimonio en los informes de la prensa, como si fuera un virus de la gripe, fue la versión mutada usada en el epígrafe de esta sección; como respuesta al consiguiente malestar, Wilson se vio forzado a vender sus acciones para conseguir la nominación y fue acosado por la «frase» durante el resto de sus días.

Podemos utilizar este accidente históricamente congelado para darle un nuevo uso. Existen pocas dudas de *por qué* la versión mutada de la frase de Wilson se propagó. Antes de que los senadores aprobaran a Charles Wilson para este importante cargo, deseaban asegurarse de quién sería el *principal beneficiario* de sus decisiones, si el país o la General Motors. ¿Iba Wilson a tomar decisiones en provecho propio o para el beneficio de la totalidad del cuerpo político? Su respuesta real dio poca seguridad a los senadores. Éstos pensaron que allí había gato encerrado y lo manifestaron en el cambio de sus palabras que dieron a conocer a todo el mundo. Wilson parecía afirmar que nadie debía preocuparse por sus tomas de decisiones, dado que si el principal o directo beneficiario era la General Motors, sería bueno para la totalidad del país. Una afirmación demasiado dudosa como para estar seguros. Aunque pudiera ser verdad la mayoría de las veces —«si todo seguía igual»—, ¿qué sucedería si cambiaban las condiciones? ¿Qué beneficio obtendría Wilson en aquellas circunstancias? Esto era lo que disgustaba a los senadores y, desde luego, con razón. Deseaban que las decisiones del secretario de Defensa beneficiaran directamente a los intereses nacionales. Si las decisiones adoptadas bajo aquellas favorables circunstancias beneficiaban a la General

Motors (y presumiblemente a la mayoría de ellos, si la homilía de Wilson era verdad) esto sería estupendo, pero los senadores temían que Wilson tuviese escondidas sus prioridades.

Este es un ejemplo de un problema de preocupación perenne y propiamente humana. Los abogados preguntan, en latín, *Cui bono?* ('¿A quién beneficia?'), una pregunta que a menudo golpea en el corazón de las cuestiones importantes: ¿a quién beneficia este asunto? El mismo problema surge en la teoría de la evolución, en la que la contrapartida de la frase real de Wilson sería: «Lo que es bueno para el cuerpo es bueno para los genes y viceversa». En términos generales, los biólogos estarían de acuerdo en que esto puede ser verdad. La suerte de un cuerpo y la de sus genes están estrechamente ligadas. Pero no son perfectamente coincidentes. ¿Qué hay de aquellos casos en los que los intereses del cuerpo (larga vida, felicidad, confort, etc.) entran en conflicto con los intereses de los genes?

Esta cuestión ha estado siempre latente en la síntesis moderna del darwinismo. Una vez que los genes han sido identificados como aquellas cosas cuya reproducción diferencial era responsable de los cambios de diseño en la biosfera, la cuestión era inevitable, aunque durante largo tiempo los teóricos podían tranquilizarse, como Charles Wilson, con la reflexión de que, en general, lo que era bueno para la totalidad era bueno para una parte y viceversa. Pero cuando George Williams [1966] llamó la atención sobre la cuestión, y la gente comenzó a darse cuenta de que había profundas implicaciones para nuestra comprensión de la evolución, Dawkins [1986] convirtió en ineludible la cuestión al enmarcarla en el concepto del gen egoísta, puntualizando que desde el «punto de vista» del gen, un cuerpo era una especie de máquina de supervivencia creada para incrementar las posibilidades de los genes de una reproducción continuada.

El viejo panglossianismo ha pensado confusamente que la adaptación era «para el bien de las especies»; Williams, Maynard Smith, Dawkins y otros demostraron que «el bien del organismo» era precisamente una visión tan miope como el «bien de las especies». Para comprobarlo, se ha de adoptar una perspectiva todavía menos engañosa, la perspectiva del gen, y preguntarse qué es lo bueno para los genes. A primera vista esta postura parece dura, fría y despiadada. Me recuerda, de hecho, aquella trillada regla empírica que se hizo famosa en las novelas de misterio: *Cherchez la femme! ¡Buscad a la mujer!*[118] La idea es que, como cualquier detective realista y conocedor del mundo sabe, la llave que abre la puerta del misterio implica a una u otra mujer, de algún modo. Probablemente es un mal consejo, incluso en el

estilizado e irreal mundo de la novela policíaca. Los que se centran en el gen sostienen un concepto mejor: *Cherchez le gène!* Ya vimos un buen ejemplo de esta postura en el relato, en el capítulo 9, de la investigación de David Haig, pero hay cientos o miles de otros ejemplos que pudieran citarse. (Cronin 1991 y la revisión de Matt Ridley 1993 sobre la historia de esta investigación hasta la actualidad). Dondequiera que encontremos un rompecabezas en la evolución, la perspectiva de la visión del propio gen está en condiciones de generar una solución en relación con un gen u otro, que resulta favorecido por una u otra razón. En cuanto a las adaptaciones, son claramente para el bien del organismo (el águila, en tanto que organismo, seguramente se beneficia de su ojo de águila y de su ala de águila); esto es así fundamentalmente por la razón expuesta por Wilson: lo que es bueno para los genes es bueno para el conjunto del organismo. Pero, a la hora de la verdad, lo que es bueno para los genes determina lo que en el futuro se ha de mantener. Los genes son, después de todo, los replicadores cuyas variables perspectivas en las competiciones de autorreplicación ponen en marcha el proceso total de la evolución, y lo mantienen en movimiento.

Esta perspectiva desde los genes, a veces denominada «gen-centrismo» o el «punto de vista desde el ojo del gen», ha provocado muchas críticas, y no pocas desencaminadas. Por ejemplo, se dice a menudo que el gen-centrismo es «reduccionista». Lo es en el buen sentido. Es decir, descarta los ganchos celestes e insiste en que todas las elevaciones en el espacio de diseño realizarse mediante grúas. Pero, como hemos visto en el capítulo 3, a veces se usa la palabra «reduccionismo» para referirse al punto de vista que sostiene que se debe «reducir» toda la ciencia o todas las explicaciones, al nivel más bajo: el nivel molecular o el atómico o el subatómico (pero probablemente nadie aún ha adoptado esta variedad de reduccionismo, que es evidentemente estúpida). En cualquier acontecimiento el gen-centrismo es manifiestamente *no* reduccionista, en este sentido del término. ¿Qué puede ser menos reduccionista que explicar la presencia de una molécula particular de un aminoácido en una localización particular en un cuerpo particular, no por algunos hechos a nivel molecular sino más bien por el hecho de que el cuerpo en cuestión era una hembra de una especie que presta cuidados maternales durante largo tiempo a su progenie? El punto de vista del ojo del gen explica cosas en términos que resaltan la intrincada relación entre hechos ecológicos de amplia variedad y a gran escala, hechos históricos a largo plazo, y hechos a nivel molecular.

La selección natural no es una fuerza que «actúe» a un nivel, por ejemplo, el nivel molecular como opuesto al nivel de la población o al nivel del organismo. La selección natural ocurre debido a que una suma de acontecimientos de todo tipo y tamaño tiene un resultado concreto estadísticamente descriptible. La ballena azul se desliza hacia el precipicio de la extinción; si se extinguiera, una serie de volúmenes de la librería de Mendel, especialmente magníficos y casi imposibles de reemplazar, dejaría de tener copias vivas, ¡pero el factor que mejor puede explicar por qué aquellos característicos cromosomas de la ballena azul, o las colecciones de secuencias de nucleótidos en su ADN, desaparecerían de la tierra, puede ser un virus que, de alguna manera, ataque directamente la maquinaria de replicación del ADN en las ballenas, un cometa extraviado que aterrizase cerca de la manada de supervivientes justamente en el momento equivocado, o un exceso de publicidad televisiva que ocasionara que los curiosos humanos interfirieran catastróficamente en su hábitos de reproducción! Hay siempre una descripción de todo hecho evolutivo desde el punto de vista del ojo del gen, pero la cuestión más importante es si tal descripción puede a menudo ser mera «contabilidad» (*bookkeeping*) (y tan poco clarificador como un marcador de béisbol a nivel molecular). William Wimsatt [1980] introdujo el término *bookkeeping* para referirse al hecho, aceptado por todos, de que los genes son el almacén de la información de los cambios genéticos, dejando abierto el debate sobre si el punto de vista centrado en el gen era solamente simple «contabilidad», una acusación que, a menudo, le ha sido imputada (por ejemplo, Gould 1992a). George Williams acepta la etiqueta aunque defiende con vigor la importancia de la contabilidad. «La idea que el *bookkeeping* ha ido adquiriendo en el pasado es lo que da a la teoría de la selección natural su más importante poder de predicción» (1985:4). (Véase Buss 1987, especialmente pp. 174 y ss., con importantes reflexiones sobre esta cuestión).

La idea de que la perspectiva del gen-centrismo es mejor o más importante no es una afirmación acerca de la importancia de la biología molecular, sino sobre algo más abstracto: sobre qué nivel hace el trabajo de explicación en la mayoría de los casos. Los filósofos de la biología han prestado más atención y han realizado contribuciones más sustanciales al análisis de este tema que a ningún otro en la teoría evolutiva. He mencionado a Wimsatt y hay otros; para escoger algunos de los mejores, David Hull [1980], Elliot Sober [1981a] y Kim Sterelny y Phil Kitcher [1988]. Una razón por la que los filósofos se han visto atraídos por esta cuestión es seguramente su carácter abstracto y lo intrincado de su concepto. Pensando en ello pronto

se entra en cuestiones profundas acerca de lo que significa explicar algo, qué es la causalidad, a qué nivel, y así sucesivamente. Ésta es una de las áreas más brillantes en la reciente filosofía de la ciencia; los científicos han prestado atención respetuosa a sus colegas filósofos y han visto esta atención recompensada con análisis eruditos, bien informados, y argumentos elaborados por los filósofos, a los cuales los científicos, a su vez, han respondido con sus propias discusiones, por encima del significado filosófico habitual. Ésta es una rica cosecha, por lo que me resulta difícil alejarme de ella sin proporcionar un introducción a las sutilezas de estos temas, pues sostengo sólidas opiniones en torno a la sabiduría que subyace en estas controversias; pero mi plan de trabajo, en este momento, es otro y consiste, curiosamente, en *desdramatizar* esta cuestión. Se trata de problemas científicos y filosóficos excelentes, pero se obtengan los resultados que se obtengan, no tendrán sobre el darwinismo el impacto que algunos temieran (este será un tema de discusión en el capítulo 16).

Las repeticiones y reflexiones que tratan de presentar una explicación evolucionista son un suplicio de Tántalo y, por ello, razón suficiente para que los filósofos presten atención a la controversia de las unidades de selección, aunque una razón adicional para esta mayor atención de los filósofos es, seguramente, la reflexión con la cual he comenzado esta sección: la gente se siente amenazada por la perspectiva que ofrece el punto de vista del gen, por la misma razón que se sentía amenazada por la fidelidad de Charles Wilson a la General Motors. La gente desea estar al frente de su propio destino; quieren ser, al mismo tiempo, los que deciden y los principales beneficiarios de sus decisiones y muchos temen que el darwinismo, en esta versión centrada en los genes [gen-centrismo], debilite su apuesta por el dominio del propio destino. Se dan cuenta de que la atrayente pintura de Dawkins, en la que el organismo aparece como un mero vehículo creado para transportar un conjunto de genes a futuros vehículos, es un asalto y una agresión intelectual. Así que, en mi opinión, una razón por la que las perspectivas planteadas a nivel de organismo y de grupo son ya bien recibidas frecuentemente como merecidos oponentes de la perspectiva desarrollada a nivel de gen es el pensamiento básico, aunque nunca articulado, de que ¡*nosotros* somos organismos (y vivimos en grupos que nos importan a nosotros) y no deseamos que nuestros intereses sean manipulados por otros! En otras palabras, mi presentimiento es que a nosotros no debe importarnos si los pinos o los colibrís son «meras máquinas de supervivencia» para sus genes, ya que no mantenemos con nuestros genes la misma relación que ellos mantienen con los suyos. En el próximo capítulo

quiero que nos quitemos de encima esta preocupación demostrando que realmente no hay motivo alguno, ya que nuestra relación con nuestros genes es muy diferente de las relaciones de otras especies con sus genes, debido a que lo que *nosotros* somos no es precisamente lo que como especies somos. De este modo, como si se retirara un tapón, podrá drenarse toda la ansiedad provocada por este fascinante problema conceptual, aún no resuelto, es decir, cómo debemos entender a los genes en tanto que unidades de selección, pero antes de que volvamos a la tarea, debemos estar seguros de que el aspecto más amenazante del tema se ha clarificado y que también se han aclarado varias concepciones falsas.

Quizá la crítica más descarriada del gen-centrismo es la afirmación, frecuentemente escuchada, de que los genes, sencillamente, no pueden tener intereses (Midgley 1979, 1983, Stove 1992.). Esta crítica, si se toma en serio, conduciría a descartar un tesoro de sugerencias intelectuales, pero es totalmente errónea. Aunque si los genes no pudieran *actuar* a favor de sus intereses, del modo que lo hacemos nosotros, pueden seguramente tenerlos en un sentido que es claro e incontrovertible. Si un cuerpo político, o la General Motors, puede tener intereses, del mismo modo lo pueden tener los genes. Nosotros podemos hacer algo por nuestro propio bien, o por el bien de los hijos, o por el bien del arte, o por el bien de la democracia, o por el bien de… de la mantequilla de cacahuete. Es difícil imaginar la razón por la que alguien podría desear colocar el bienestar o incluso el florecimiento de la mantequilla de cacahuete por encima de todo eso, pero la mantequilla de cacahuete puede estar en el pedestal del mismo modo que pueden estarlo el arte o los hijos. Uno puede incluso decidir —aunque sería algo extraño— que la cosa que desea proteger y potenciar aun a costa de su propia vida, fueran los propios genes. Ninguna *persona* en sus cabales tomaría esta decisión. Como dice George Williams: «No existe una justificación concebible para una preocupación personal por los intereses (proliferación media a largo plazo) de los genes que hemos recibido en la lotería de la meiosis y la fertilización» (1988:403).

Pero esto no significa que no haya *fuerzas* que tienden a promover el bien o los intereses de los genes. En efecto, hasta hace muy poco, los genes eran los principales beneficiarios de todas las fuerzas selectivas del planeta. Es decir, no había fuerzas cuyos *principales* beneficiarios fueron otros. Hubo *accidentes* y *catástrofes* (rayos y maremotos), pero no fuerzas *que de manera estable* actuasen sistemáticamente en favor de algo que no fueran los genes.

¿Hacia qué intereses se orienta, de forma más directa, el proceso real de la toma de decisiones de la selección natural? No es motivo de controversia que pueden surgir conflictos entre genes y cuerpos (entre genes y la expresión fenotípica de los genotipos de los cuales forman parte). Por otro lado, nadie duda de que, en general, las exigencias del cuerpo de ser considerado el principal beneficiario, se suspenden tan pronto como ha completado su misión procreadora. Una vez que el salmón ha luchado contracorriente subiendo aguas arriba y desovando con éxito, se transforma en carne muerta. Literalmente se cae a pedazos, debido a que ya no existe una presión evolutiva en favor de alguna de las revisiones del diseño que puedan prevenir que los salmones desaparezcan, permitiéndoles largos y agradables períodos de retiro como abuelos, tal como muchos de nosotros llegamos a disfrutar. En general, el cuerpo es entonces solamente un instrumento y, en consecuencia, algo secundario, beneficiario de las «decisiones» tomadas por la selección natural.

Esto es verdad a través de la biosfera, revelado en un modelo con unas pocas e importantes variaciones. En muchas *phyla*, los padres mueren antes de que sus vástagos hayan nacido y la totalidad de sus vidas son una preparación para el clímax de un simple acto de replicación. Otros —los árboles, por ejemplo— viven a través de muchas generaciones de vástagos y pueden, en consecuencia, competir por la luz solar y por otros recursos con su propia progenie. Los mamíferos y las aves invierten, de modo característico, grandes cantidades de energía y de actividad en el cuidado de su progenie y, en consecuencia, tienen muchas oportunidades para «escoger» entre ellos mismos y su progenie como beneficiarios de cualquier acción que asumen. Las criaturas a las que tales opciones nunca llegan pueden estar diseñadas «bajo la premisa» (una tácita premisa de la madre naturaleza) de que éste no es, sencillamente, un tema que necesite ninguna atención al diseño.

Cabe presumir que el sistema de control de una polilla, por ejemplo, está despiadadamente diseñado para sacrificar el cuerpo por el bien de los genes, dondequiera que pueda surgir una oportunidad genérica y reconocible. Una pequeña fantasía: nosotros, de alguna manera, reemplazamos este sistema estándar en la biosfera (el sistema que se expresaría diciendo: «¡malditos torpedos!, ¡a toda velocidad hacia delante!») por un sistema que favorezca al cuerpo (el sistema que se expresaría con la decisión: «¡al infierno con mis genes! ¡Estoy preparado para ser el número uno!»). ¿Podría el reemplazamiento incluso hacer que no hubiera una vía u otra para cometer suicidio o vagar sin rumbo? Una polilla no está simplemente equipada para

conseguir ventajas de oportunidades tangenciales en el trabajo de su vida, que es reproducirse a sí misma. Es difícil tomar en serio la potenciación de la vida como fin, si se trata de la corta vida de una polilla. Las aves, por el contrario, pueden abandonar un nido lleno de huevos cuando se encuentran amenazadas, de un modo o de otro, y esto se parece más a lo que nosotros hacemos a menudo, pero la razón de que las aves puedan hacer esto es que tienen la posibilidad de iniciar la construcción de otro nido, si no en esa estación, en la próxima. Están ahora preparadas para ser el número uno, aunque sólo sea porque le dan a sus genes una oportunidad mejor de replicarse más tarde.

Nosotros somos diferentes. Hay un enorme campo de alternativas en la vida humana, pero la cuestión es la siguiente: ¿cómo y cuándo se estableció este campo? No puede haber duda de que mucha gente ha *escogido* con claridad mental y buena información seguir los riesgos y dolores de la gestación frente a la seguridad y comodidad de una vida «yerma» de otras recompensas. La cultura puede jugar contra esto (con palabras con tanta carga como «yerma») y es cierto que se trata de una inversión de la fundamental estrategia de toda vida, pero sucede a menudo. Reconocemos que tener y cuidar a la progenie es uno de los posibles proyectos vitales y, sin duda, el más importante, dados nuestros valores. Pero ¿de dónde vienen estos valores? ¿Cómo se equipó con ellos nuestro sistema de control, si no fue mediante una intervención quirúrgica milagrosa? ¿Cómo ha sucedido que hemos llegado a ser capaces de establecer una perspectiva rival que puede, a menudo, sobrepasar los intereses de nuestros genes mientras que otras especies no pueden[119]? Este será el tema del próximo capítulo.

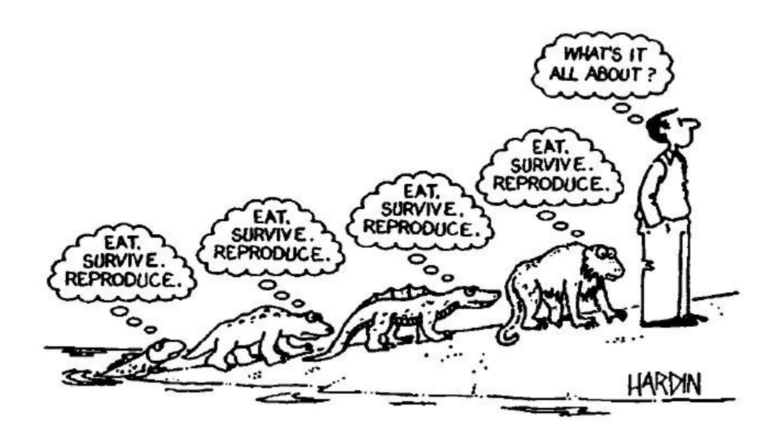


Figura 11.1

PARTE III
Mente, significado, matemáticas y moralidad

El nuevo sentimiento fundamental: nuestra definitiva transitariedad. En un principio se buscó el sentimiento de la grandeza del hombre señalando su origen divino: éste es ahora un camino vedado, ya que a su entrada se encuentra el mono, ¡unto con otras repugnantes bestias, haciéndonos muecas, como si nos quisieran decir: ¡no sigáis en esta dirección! En consecuencia, se prueba en la dirección opuesta: la vía que está siguiendo la humanidad servirá como prueba de su grandeza y de su parentesco con Dios. ¡Qué le vamos a hacer! ¡Este camino también es vano! Al término de su recorrido se encuentra la urna funeraria del *último* hombre y sepulturero (con la inscripción *nihil humani a me alienum puto*). Cualquiera que sea la altura a la que pueda haber llegado la humanidad en su evolución —¡quizás al final podría estar más abajo de lo que estaba al principio!— no cabe pasar a un orden más elevado, del mismo modo que la hormiga y la tijereta no pueden, al final de su «tránsito terrestre», elevarse al parentesco con Dios y a la vida eterna. El devenir arrastra tras de sí lo que ha sido: ¿por qué se debe hacer una excepción en este eterno espectáculo para beneficio de algún pequeño astro o de cualquiera pequeña especie? ¡Olvidemos estos sentimentalismos!

Friedrich Nietzsche, *Aurora*

CAPÍTULO 12
Las grúas de la cultura

I. *El tío del mono encuentra el «meme»*

> ¿Cuál es la cuestión a la que ahora se enfrenta la sociedad con una seguridad poco convincente, realmente asombrosa? La cuestión es ésta: ¿Es el hombre un mono o un ángel? ¡Dios mío!, estoy del lado de los ángeles.
>
> Benjamin Disraeli
> (*Discurso en Oxford*, 1864)

Hasta Darwin comprendió lúcidamente que si afirmaba que su teoría era aplicable a una especie en particular, disgustaría a los miembros de esta especie en el sentido que temía, así que, al principio, se contuvo. En *El origen de las especies*, casi no se menciona a nuestra especie, fuera de su importante papel como grúa en la selección artificial. Pero, naturalmente, esta actitud no engaña a nadie. Estaba claro cuál era el objetivo de la teoría, así como que Darwin había trabajado duramente para producir su propia versión, cuidadosamente elaborada, antes de que los críticos y los escépticos pudieran enterrarla con falsas lecturas y gritos de alarma: *El origen del hombre* (1871). No hay duda en absoluto de la posición que Darwin expresó diciendo: nosotros —*Homo sapiens*— somos una de las especies sobre las que reina la teoría de la evolución. Viendo que existía poca esperanza de que pudiera negarse este hecho, algunos de los que temían a Darwin buscaron a un defensor que pudiera poner en marcha un ataque preventivo, incapacitando la peligrosa idea antes de que pudiera extenderse a través del istmo que conecta a nuestra especie con todas las demás. Dondequiera que encontraron a alguien que anunciara la abdicación del darwinismo (o el neodarwinismo o la síntesis moderna del darwinismo) lo estimularon con la esperanza de que esta vez la revolución sería una realidad. Estos peculiares revolucionarios golpearon al darwinismo desde muy pronto y repetidas veces aunque, como hemos visto, lo único que han conseguido ha sido fortalecer su objetivo, profundizando

nuestra comprensión del darwinismo, al tiempo que lo enriquecían con complejidades que Darwin ni siquiera soñó.

Batiéndose en retirada, algunos de los enemigos de la peligrosa idea de Darwin se plantaron con firmeza en el istmo, como Horacio en el puente, en un intento de impedir que la idea lo cruzara. El primer enfrentamiento tuvo lugar en el famoso debate celebrado el año 1860 en el Museo de Historia Natural de Oxford, sólo unos meses después de la primera publicación de *El origen de las especies*, entre el obispo de Oxford, Samuel Wilberforce, conocido como el «jabonoso Sam» por su habilidad para escabullirse del contrario en los debates, y Thomas Henry Huxley, apodado «el bulldog de Darwin» (debido a su feroz defensa de la idea darwiniana). Esta es una historia mil veces contada y con tantas variantes que podemos encontrar en ella no precisamente una especie, sino un *phylum* de «memes». Aquí se encuentra el famoso error retórico del buen obispo, al preguntar a Huxley si descendía de un mono por parte de su abuelo o por parte de su abuela. La irritación alcanzó niveles elevados en la sala de reuniones; una dama se desmayó y varios de los defensores de Darwin estuvieron a punto de perder los estribos, furiosos ante la despectiva y falsa presentación de la teoría de su héroe, por lo que es comprensible que las versiones de los testigos difieran en este punto. En la mejor versión —la cual, con toda probabilidad, ha sufrido cambios significativos para mejorar su diseño a lo largo de las veces que ha sido contada y recontada— Huxley replicó que él «no se avergonzaba de tener como antepasado a un mono, pero que sí se sentiría avergonzado de tener relación con un hombre que utilizaba sus grandes facultades para oscurecer la verdad» (R. Richards 1987:4; véanse también 549-551, y Desmond y Moore 1991 reap. 33).

Desde entonces, algunos miembros del *Homo sapiens* se han mostrado excesivamente susceptibles ante nuestras ancestrales relaciones con los simios. Cuando Jared Diamond publicó *El tercer chimpancé* en 1992, se inspiró para su título en el reciente descubrimiento hecho de que los seres humanos estamos, en realidad, más relacionados con las dos especies de chimpancés (el *Pan troglodytes*, la más frecuente, y el *Pan paniscus*, el raro y pequeño chimpancé pigmeo o bonobo) que lo que estos chimpancés lo están con los otros simios. Nuestras tres especies tienen un ancestro común más reciente que el ancestro común del chimpancé y del gorila, por ejemplo, así que estamos todos sobre una rama del árbol de la vida, con gorilas y orangutanes, y los demás sobre otras ramas.

Somos el tercer chimpancé. Diamond extrajo este fascinante hecho, con prudencia, del trabajo «filológico» realizado por Sibley y Ahlquist [1984] sobre el ADN de los primates y dejó claro a sus lectores que los suyos eran un conjunto de estudios, en cierro modo controvertidos (Diamond 1992: 20, 371-372). Sin embargo, Diamond no fue lo suficientemente prudente, en opinión de otro crítico de su trabajo. Jonathan Marks, un antropólogo de Yale, se colocó en la órbita de la denuncia de Diamond, así como de Sibley y Ahlquist, cuyas publicaciones, según él, «necesitan tratarse como los desechos nucleares: enterradas con seguridad y olvidadas durante un millón de años» (Marks 1993a:61). Desde 1988, Marks, cuyas propias investigaciones, anteriores a las de Sibley y Ahlquist, habían colocado al chimpancé en una posición *marginal*, más cercana del gorila que de nosotros, se ha empeñado en una campaña alarmantemente llena de vituperios, condenando a Sibley y a Ahlquist, aunque esta campaña ha sufrido recientemente un tropiezo. Los hallazgos originales de Sibley y Ahlquist han sido rotundamente confirmados mediante métodos analíticos más sensibles (los suyos eran relativamente burdos, pioneros en su tiempo, pero posteriormente sobrepasados por técnicas más poderosas). Con todo, ¿por qué debe establecerse una diferencia moral entre nosotros o los gorilas para ver quién ha ganado la competición, para ver quién es el primo más cercano de los chimpancés? Los simios son, en cualquier caso, nuestros parientes más cercanos. Pero según parece esto preocupa muchísimo a Marks, cuyos deseos de desacreditar a Sibley y a Ahlquist lo han sacado de sus casillas. Su ataque más reciente, en una revisión de otros libros en la revista *American Science* (Marks 1993b), presenta un coro de colegas científicos con sus condenaciones, y, además, una notable apología de los editores de la revista que dice: «Aunque las opiniones de los críticos de libros son personales y no de la revista, los editores exigen unas normas por lo que sentimos profundamente que no hayan sido respetadas en la evaluación en cuestión» (septiembre-octubre de 1993:407). Como le sucedió antes al obispo Wilberforce, Marks perdió los estribos.

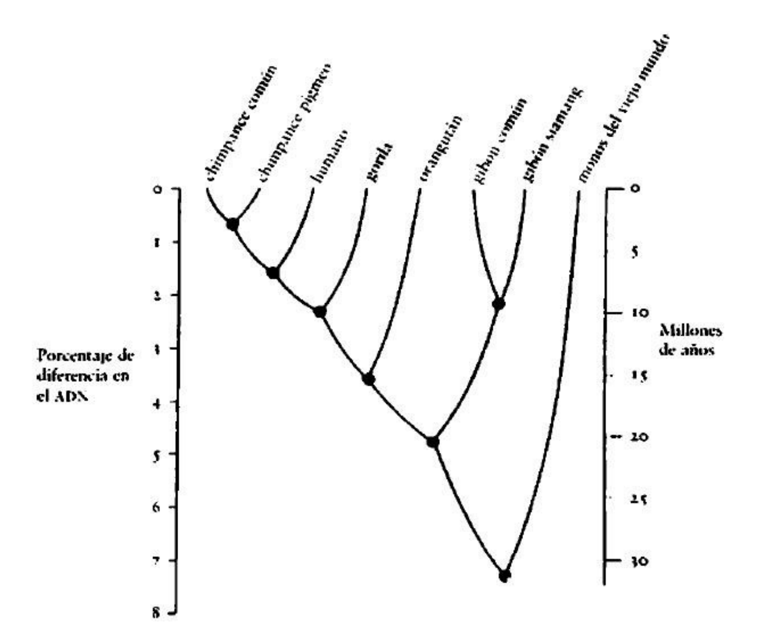


Figura 12.1

El árbol de la familia de los primates superiores en el que se remonta la filiación de cada par de primates modernos hasta el punto negro en donde se conectan. Los números a la izquierda representan la diferencia en porcentaje entre los ADN de estos primates modernos, mientras que los números a la derecha representan el número estimado en millones de años desde que compartieron por última vez un antepasado común. Por ejemplo, los chimpancés común y pigmeo difieren aproximadamente en el 0,7 por ciento de ADN se separaron hace unos tres millones de años; nosotros diferimos en el 1,6 por ciento de nuestro ADN de ambos chimpancés y nos separamos de nuestro antepasado común hace unos siete millones de años; los gorilas difieren en alrededor de un 2,3 por ciento de su ADN de nosotros o de los chimpancés y se separaron del antepasado común del que procedemos nosotros y los chimpancés hace aproximadamente diez millones de años (Diamond 1992).

La gente quiere creer que nosotros, los seres humanos, somos enormemente diferentes del resto de las especies y ¡están en lo cierto! Somos diferentes. Somos la única especie que posee un medio *extra* para preservar su diseño y comunicarlo: la cultura. Esta es una declaración excesiva; otras especies tienen rudimentos de cultura y su capacidad para transmitir información mediante la conducta, en adición a la transmitida genéticamente, es, en sí

misma, un importante fenómeno biológico (Bonner 1980), aunque estas otras especies no han desarrollado cultura hasta el punto de despegue alcanzado por nuestra especie. Tenemos lenguaje, el medio primario de la cultura, y el lenguaje ha abierto nuevas regiones en el espacio de diseño, cuyo secreto solamente nosotros poseemos. En unos pocos milenios —un mero instante en el tiempo biológico— hemos utilizado nuestros nuevos vehículos de exploración para transformar no sólo nuestro planeta, sino el propio proceso de desarrollo del diseño que nos ha creado.

La cultura humana, como hemos visto, no es precisamente una grúa compuesta de grúas, sino una grúa que hace grúas. La cultura es un conjunto de grúas tan poderoso que sus efectos pueden frenar muchas —si no todas— de las más tempranas presiones genéticas y de los procesos que la han creado y aún coexisten con ella. A menudo cometemos el error de confundir una innovación cultural con una innovación genética. Por ejemplo, todo el mundo sabe que la altura media de los seres humanos se ha disparado en los últimos siglos. (Cuando visitamos en el puerto de Boston reliquias de la reciente historia como el viejo *Old Ironsides*, el buque de guerra de inicios del siglo XIX, observamos que el espacio bajo cubierta es casi cómicamente angosto; ¿eran nuestros antepasados una raza de enanos?). ¿Hasta qué punto este rápido cambio de estatura se debe a cambios genéticos en nuestra especie? Poco, o nada. ¿Ha habido tiempo en tan sólo diez generaciones de *Homo sapiens* desde que el *Old Ironsides* fuera botado en 1797, Y aunque hubiera habido una fuerte presión de selección para favorecer a los más altos, hay evidencia de ello? No ha habido tiempo para que se produzca tan gran efecto. Lo que ha cambiado extraordinariamente ha sido la salud humana, la dieta y las condiciones de vida; éstas son las que han producido este enorme cambio en el fenotipo, el cual se debe cien por cien a innovaciones culturales, pasadas a través de transmisiones culturales: escolarización, nuevas prácticas de cultivo, medidas de salud pública, etc. A cualquiera que le preocupe el «determinismo genético» se le debe recordar que virtualmente todas las diferencias discernibles en la gente, desde los días de Platón hasta nuestro tiempo, sus talentos físicos, proclividades, actitudes, perspectivas, se deben a cambios culturales, dado que poco más de doscientas generaciones nos separan del filósofo griego. Sin embargo, los cambios ambientales debidos a innovaciones culturales cambian el paisaje de la expresión fenotípica, tanto y a tanta velocidad que pueden cambiar rápidamente, en principio, la presión de selección genotípica; el efecto Baldwin es un simple ejemplo de este cambio en la presión de selección debido a la amplia innovación de la conducta.

Mientras que es importante recordar cuán lentamente trabaja la evolución, en general, no debemos nunca olvidar que no hay en absoluto inercia en la presión de selección. Presiones que han sido dominantes durante millones de años pueden desvanecerse en un santiamén; naturalmente, nuevas presiones de selección pueden aparecer con una simple erupción volcánica, o con la aparición del organismo de una nueva enfermedad.

La evolución cultural opera con mucha mayor rapidez, en muchos órdenes de magnitud, que la evolución genética, y este es parte de su papel en la función de hacer de nuestra especie algo especial, aunque también ha dado como resultado el convertirnos en criaturas con una perspectiva de la vida completamente diferente de las otras especies. En efecto, no está claro que los miembros de cualquier otra especie que no sea la humana tengan una perspectiva de la vida. Pero nosotros sí la tenemos; podemos escoger el celibato por varias razones; podemos aprobar leyes que regulen lo que comemos; podemos tener sistemas elaborados para estimular o castigar ciertos tipos de conducta sexual, etc. Nuestra visión de la vida es tan apremiante y obvia para nosotros que, a menudo, caemos en la trampa de imponerla, quieran o no, a otras criaturas o incluso a toda la naturaleza. Uno de mis ejemplos favoritos de esta extendida ilusión cognitiva es el rompecabezas en que se han convertido las explicaciones que sobre el sueño vienen dando los investigadores:

> Los estantes de los laboratorios se hunden bajo el peso de volúmenes de datos, aunque todavía nadie ha discernido que el sueño tiene una clara función biológica. ¿Qué presión evolucionista seleccionó esta curiosa conducta que nos obliga a permanecer inconscientes un tercio de nuestra vidas? Los animales dormidos son más vulnerables frente a los predadores. Tienen menos tiempo para buscar alimentos, para comer, para encontrar pareja, para procrear, para alimentar a sus hijos. Como decían a sus hijos los padres Victorianos, los dormilones se quedan atrás en la carrera de la vida y en la evolución.
>
> Allan Rechsthaff, investigador de la Universidad de Chicago, se pregunta: «¿cómo pudo la selección natural “permitir”, con su irrevocable lógica, al reino animal pagar el precio del sueño por razones que no son buenas?»». El sueño es tan claramente una mala adaptación que es difícil comprender por qué no ha evolucionado alguna otra condición para satisfacer las necesidades que el sueño satisface (Raymo 1988).

Pero ¿por qué necesita el sueño «una clara función biológica»? Es el *estar despierto* lo que necesita una explicación y, posiblemente, su explicación es obvia. Los animales —al contrario que las plantas— necesitan estar despiertos, al menos una parte de su tiempo vital, para buscar alimento y procrear, como recuerda Raymo. Pero una vez que hemos escogido el camino de llevar una existencia activa, el análisis coste-beneficio de las opciones que surgen está lejos de ser obvio. Estar despierto es relativamente costoso, comparado con yacer dormido. De este modo, presumiblemente, la madre

naturaleza economiza donde puede. Si pudiéramos seguir por este camino, estaríamos dormidos toda la vida. Esto es lo que hacen, después de todo, los árboles: durante todo el invierno «hibernan» en coma profundo, porque no hay nada que hacer y en el verano «veranean» en un estado de coma algo más ligero, lo que los médicos llaman *estado vegetativo*, lo que sucede cuando un miembro de nuestra especie tiene la mala fortuna de caer en ese estado. Si el leñador llega cuando los árboles están durmiendo, bien, esta es justamente la posibilidad que, en todo momento, los árboles han de tener en cuenta. No obstante, ¿corremos nosotros, los animales, un riesgo mucho mayor mientras dormimos? Abandonar la madriguera es también arriesgado, y si intentamos minimizar esta fase de riesgo, también podemos mantener el metabolismo inactivo mientras aguardamos el momento apropiado, conservando energía para el asunto principal que nos ocupa, que es la replicación.

(Desde luego, estos asuntos son bastante más complicados de como los presento. Mi posición es que el análisis de coste-beneficio no es nada obvio y que es suficiente con eliminar el aire de paradoja).

Nosotros creemos que ir de un lado para otro, corriendo aventuras y completando proyectos, visitando a nuestros amigos y aprendiendo a conducirnos en el mundo, es la cuestión fundamental de la vida, pero la madre naturaleza no ve las cosas de la misma manera. Una vida de sueño es una vida tan buena como cualquier otra, y en muchos aspectos mejor —ciertamente más barata— que la mayoría. Si los miembros de otras especies también parecen disfrutar de sus períodos en estado de alerta como hacemos nosotros, ésta es una interesante característica compartida, tan interesante que no debemos cometer el error de asumir que existe porque nosotros la consideramos una actitud apropiada en relación con nuestra vida. Su existencia en otras especies necesita ser demostrada, y esto no es fácil[120].

Lo que somos depende, en gran parte, de la cultura que nos ha hecho. Cabe preguntarse ahora cómo comenzó todo esto. ¿Qué tipo de revolución evolutiva sucedió que nos apartó tan decisivamente de otros productos de la evolución genética? Voy a volver a contar el relato que encontramos en el capítulo 4, a propósito de la creación de las células eucariotas, las que han hecho posible la vida multicelular. Recordemos que antes de que hubiera células con núcleos existían unas formas de vida más simples, y más solitarias, las procariotas, destinadas para algo tan simplemente imaginativo como deslizarse dentro de una especie de caldo rico en energía y allí reproducirse. No es que sea nada, pero no es lo bastante para una vida. Entonces, un día, según el maravilloso relato de Lynn Margulis [1981],

algunas procariotas fueron invadidas por parásitos de una clase inferior, y esta invasión se convirtió en una bendición encubierta, porque, mientras que los parásitos son —por definición— deletéreos en la capacidad de adaptación de sus huéspedes, estos invasores se transformaron en beneficiosos y, en consecuencia, se convirtieron en *simbiontes*, dejando de ser parásitos. Ellos y aquellos a los que invadieron, llegaron a ser más bien *comensales* —literalmente, derivado del latín, organismos que comen en la misma mesa— o *mutualistas*, en el sentido de que cada uno se beneficia de la compañía del otro. Unieron fuerzas y crearon una nueva clase de entidad revolucionaria, la célula eucariota. Con esta conjunción se abrió el vasto espacio de posibilidades que conocemos como vida multicelular, un espacio previamente inimaginable, como mínimo; las células procariotas, sin duda, desconocen estas cuestiones.

Pasaron entonces unos cuantos miles de millones de años, en tanto que las formas de vida multicelular exploraban varios rincones y grietas en el espacio de diseño hasta que un hermoso día comenzó otra invasión, esta vez en una única especie de organismo multicelular, un tipo de primate, en el que se había desarrollado una variedad de estructuras y de capacidades (no nos atrevemos a llamarlas preadaptaciones) las cuales eran especialmente adecuadas para estos invasores. No es sorprendente que los invasores estuviesen bien adaptados para encontrar alojamiento en sus huéspedes, dado que ellos mismos habían sido creados por sus huéspedes, del mismo modo que las arañas crean sus redes y los pájaros sus nidos. En un parpadeo —menos de cien mil años— estos nuevos invasores transformaron a los monos, que eran sus *inadvertidos* huéspedes, en algo nuevo: huéspedes *advertidos*, quienes, gracias a su enorme stock de recién inventados invasores, podían imaginar lo hasta ahora inimaginable, saltando a través del espacio de diseño, como nunca se había hecho anteriormente. Siguiendo a Dawkins [1976] llamo a los invasores *memes* y esta radicalmente nueva clase de entidad creada cuando un tipo particular de animal resulta apropiadamente amueblado —o infestado con *memes*—, es lo que habitualmente llamamos una *persona*.

Esta es la historia contada a grandes trazos. Me di cuenta de que algunos odian esta idea. Prefieren la idea de que nuestra mente humana —y la cultura humana— es la que nos distingue claramente de todos los «insensatos animales» (como los llamaba Descartes) aunque no les gusta la idea de tener que dar una explicación evolucionista de la creación de esta marca tan importante que nos distingue. Creo que cometen un grave error[121]. ¿Desean un milagro? ¿Desean que la cultura sea algo dado por Dios? ¿Un gancho

celeste y no una grúa? ¿Por qué? Desean que el camino humano de la vida sea radicalmente diferente de la forma de vida de los otros seres vivos y así es, aunque la cultura debe tener un origen darwiniano, como la vida misma y cualquier otra cosa maravillosa. Debe crecer a partir de algo que es menos, algo *cuasi*, algo simplemente *como si* más que algo *intrínseco* y, a cada paso a lo largo del camino, los resultados han de ser, como señala David Haig, *exigibles desde el punto de vista evolucionista*. Para la cultura necesitamos, por ejemplo, el lenguaje, pero este ha de evolucionar primero en su propio gancho; no hemos podido darnos cuenta de cuán bueno sería hasta que no lo vimos totalmente en su lugar. No podemos dar por supuesta la cooperación; no podemos dar por supuesta la *inteligencia* humana; no podemos dar por supuesta la tradición; todo esto tuvo que ser construido afanosamente, del mismo modo que los replicadores originales. Conformarse con algo menos en el camino de esta explicación sería justamente abandonar.

En el próximo capítulo, me ocuparé de importantes cuestiones teóricas acerca de cómo el lenguaje y la mente humana pudieron evolucionar, en un principio, por medio de mecanismos darwinianos. Tendremos que afrontar y desarmar la tremenda —y tremendamente descaminada— animosidad contra esta historia y también elaborar respuestas para las objeciones responsables. Pero antes de que consideremos cómo se construyó esta magnífica estructura de la grúa, quisiera dibujar esquemáticamente el producto completo, distinguirlo de sus caricaturas y mostrar, con más detalles, cómo la cultura llega a poseer tales poderes revolucionarios.

2. *La invasión de los ladrones de cadáveres*

> Los seres humanos deben su supremacía biológica a la posesión de una herencia completamente diferente de la que corresponde a los otros animales: herencia exogénica o exosomática. En esta forma de herencia, la información se transmite de una generación a la siguiente a través de canales no genéticos, por ejemplo mediante la palabra hablada, y por otras formas de adoctrinamiento; en general, por todo el aparato de la cultura.
>
> Peter Medawar, «*Unnatural Science*»

> Los ácidos nucleicos inventaron a los seres humanos con el fin de que ellos mismos pudieran reproducirse, incluso en la Luna.
>
> *Sol Spiegelman*
> (Citado en Eigen, *Steps towards Life*)

> Estoy convencido de que las comparaciones entre la evolución biológica y los cambios en la cultura y en la tecnología humanas han hecho muchísimo más daño que beneficio; abundan los ejemplos de éstas trampas intelectuales… La evolución biológica es potenciada mediante la selección natural, y la evolución cultural por una diferente serie de principios que comprendo aunque no con la suficiente claridad.
>
> Stephen Jay Gould,
> «*The Panda's Thumb of Technology*»

Nadie desea reinventar la rueda, un mítico ejemplo de un innecesario trabajo de diseño, y no tengo intención de cometer aquí este error. Hasta ahora me he ayudado del término meme propuesto por Dawkins para nombrar a cualquier ítem de evolución cultural, posponiendo la discusión de cualquier tipo de teoría darwiniana acerca de los memes. Ha llegado el momento de considerar más cuidadosamente lo que son o pueden ser los memes de Dawkins. Él mismo ha llevado a cabo una parte importante del trabajo básico de diseño (diseñando, naturalmente, sobre el trabajo previo de otros) y también yo he producido anteriormente un meme de su meme, dedicando considerable tiempo y esfuerzo a construir vehículos de explicaciones apropiadas de este tema. Voy a utilizar de nuevo estas primeras construcciones, con el añadido de posteriores modificaciones en su diseño. En primer lugar, presentaré mi propia versión (Dennett 1990c) de la teoría de Dawkins sobre los memes, que expuse en una de las «Mandel Lecture» de la American Society for Aesthetics, serie de conferencias organizadas con el propósito de analizar la cuestión de si el arte promueve la evolución humana. (¡La respuesta es Sí!). En aquella ocasión reciclé mi propio artefacto, reutilizándolo, con modificaciones, en mi libro sobre la conciencia humana (Dennett 1991a:199-208), para demostrar cómo los memes pueden transformar el sistema operativo o la arquitectura computacional del cerebro humano. Mi presentación contiene muchos detalles acerca de las relaciones entre el *hardware* del cerebro humano, diseñada genéticamente, y los hábitos transmitidos culturalmente que lo transforman en algo mucho más poderoso; pasaré de puntillas sobre la mayor parte de los detalles. En esta ocasión modificaré mi reciclamiento de Dawkins, por segunda vez, para disponer de la mejor versión para tratar los particulares problemas ambientales encontrados en el actual proyecto de explicación. (Aquellos que estén familiarizados con cualquiera de los inmediatos ancestros de esta explicación, encontrarán importantes mejoras en la versión actual).

El esquema de la teoría de la evolución mediante selección natural deja claro que la evolución ocurre donde quiera que se den las siguientes

condiciones:

1) variación: hay una continua abundancia de elementos diferentes.

2) herencia o replicación: los elementos tienen la capacidad de crear copias o réplicas de sí mismos.

3) Diferencias en la aptitud: el número de copias de un elemento creadas en un tiempo dado varía dependiendo de las interacciones entre las características de este elemento y las del medio ambiente en el cual persiste.

Nótese que esta definición, aunque extraída de la biología, no dice nada específico acerca de las moléculas, de la nutrición y ni siquiera de la vida. Esta definición de la evolución por selección natural, abstracta al máximo, ha sido formulada a grandes rasgos1 con muchas versiones equivalentes; por ejemplo, Lewontin (1980) y Brandon [1978] ambas reproducidas en Sober [1984b]. Como ha señalado Dawkins, el principio fundamental es

> que toda la vida evoluciona mediante la supervivencia diferencial de entidades replicantes…
>
> El gen, la molécula de ADN, es la entidad replicante que prevalece en nuestro propio planeta. Puede haber otras. Si las hay, siempre que se reúnan ciertas condiciones, es casi inevitable su tendencia a constituirse en la base de un proceso evolutivo.
>
> Pero ¿hemos de ir a mundos distantes para encontrar otro tipo de replicación y, en consecuencia, otras clases de evolución? Yo creo que una nueva clase de replicador ha emergido recientemente en nuestro planeta. Nos mira fijamente a la cara. Se encuentra aún en la infancia, deslizándose torpemente en su sopa primigenia, aunque ya está consiguiendo cambios evolutivos a un ritmo que deja bastante atrás, palpitando, a los viejos genes (Dawkins 1796:206).

Estos nuevos replicadores son, en cierto modo, ideas. No las «ideas simples» de Locke y de Hume (la idea de rojo o la idea de redondo o de caliente o de frío) sino ese clase de ideas complejas que forman por sí mismas *unidades memorables distintas*, como las ideas de

arco
rueda
ropa de vestir
vendetta
triángulo recto
alfabeto
calendario
la *Odisea*
cálculo
ajedrez
dibujo de perspectiva
evolución mediante selección
impresionismo
deconstruccionismo

Intuitivamente las vemos como unidades culturales más o menos identifica bles, aunque podemos decir algo más preciso acerca de cómo trazar sus límites; por qué D-F # —A no es una unidad, y sí lo es el tema del

movimiento lento de la Séptima Sinfonía de Beethoven: las unidades son los elementos más pequeños que se replican a sí mismos con fiabilidad y fecundidad. Podemos compararlas, en este sentido, con los genes y sus componentes: C-G-A, un simple codon de ADN, es «demasiado pequeño» para ser un gen. Es uno de los codones del aminoácido arginina, y se copia prodigiosamente a sí mismo donde quiera que aparece en los genomas, aunque sus efectos no son lo suficientemente «individuales» como para ser considerado un gen. Una frase de tres nucleótidos no cuenta como un gen por la misma razón que no podemos registrar la propiedad intelectual de una frase musical de tres notas: no es suficiente para hacer una melodía. Pero no hay, en principio, un límite inferior para la longitud de la secuencia que pueda llegar a ser considerada un gen o un meme (Dawkins 1982:89ss). Las primeras cuatro notas de la Quinta Sinfonía de Beethoven son, sin duda, un meme, replicándose a sí mismas, separadas del resto de la sinfonía, aunque manteniendo intacta una cierta identidad de efecto (un efecto fenotípico) y, en consecuencia, creciendo en contextos en los cuales Beethoven y sus obras eran desconocidos. Dawkins explica cómo acuñó el nombre que dio a estas unidades:

> Una unidad de transmisión cultural, o una unidad de *imitación*. «Mímeme» se deriva de una apropiada raíz griega, pero yo prefiero un monosílabo que suene un poco como «gene» (gen). El término «meme» podría alternativamente considerarse relacionado con «memoria», o con la palabra francesa même…
>
> Ejemplos de memes son tonadas, ideas, frases hechas que se ponen de moda, la moda en el vestir, la manera de hacer ollas o de construir arcos. Del mismo modo que los genes se propagan en el *pool* de genes, saltando de cuerpo en cuerpo por la vía del esperma o de los huevos, así los memes se propagan en el *pool* de los memes saltando de cerebro en cerebro a través de un proceso que, en el buen sentido de la palabra, se puede llamar *imitación*. Si un científico escucha o lee alguna buena idea, la transmite a sus colegas y estudiantes. La menciona en sus artículos y en sus conferencias. Si la idea atrae la atención, puede decirse que se propaga por sí misma, extendiéndose de cerebro en cerebro (Dawkins 1976:206).

La evolución del meme, de acuerdo con Dawkins, no es exactamente análoga a la evolución biológica o genética. No se trata de un proceso que pueda describirse metafóricamente en estos idiomas evolucionarlos, sino de un fenómeno que obedece a las leyes de la selección natural con gran exactitud. La teoría de la evolución por selección natural es neutral, sugiere Dawkins, con respecto a las diferencias entre memes y genes; se trata de diferentes clases de replicadores que evolucionan en medios diferentes a ritmos diferentes. Y precisamente, del mismo modo que los genes para los animales no pudieron llegar a la existencia en este planeta hasta que la evolución de las plantas había recorrido su camino (creando la atmósfera rica en oxígeno y el cúmulo suficiente de nutrientes convertibles), así la evolución de los memes

no pudo haber comenzado hasta que la evolución de los animales había recorrido el camino que le llevó a crear una especie —*Homo sapiens*— con cerebros capaces de proporcionar refugio y hábitos de comunicación que facilitasen a los memes medios de transmisión.

No hay que negar que existe una evolución cultural, en el sentido neutral darwiniano de que las culturas cambian a lo largo de los tiempos, acumulando y perdiendo características, aunque también mantienen características de las edades más tempranas. La historia de una idea, por ejemplo, la crucifixión, o la cúpula sobre pechinas, o el vuelo con motor, es sin duda la historia de la transmisión de una familia de variaciones sobre un tema central, a través de varios medios no genéticos. Pero es una cuestión abierta si tal evolución es análoga, débil o fuertemente, o bien paralela a la evolución genética, ese proceso que la teoría darwiniana explica tan bien. De hecho, hay muchas cuestiones abiertas. En un extremo, imaginemos que sucediera que la evolución cultural recapitulara *todas* las características de la evolución genética: no sólo existirían análogos de los genes (memes) sino análogos estrictos de fenotipos, genotipos, reproducción sexual, selección sexual, ADN, ARN, codones, especiación alopátrica, demes, impresión genómica, etc., todo el edificio de la teoría biológica perfectamente reflejado a modo de especulación en el medio de la cultura. ¿Piensa que la ingeniería genética del ADN es una tecnología que causa pavor? ¡Pues espere a que inicien implantes de memes en sus laboratorios! No parece probable. En el otro extremo, podría descubrirse que la evolución cultural funciona de acuerdo con principios totalmente diferentes (como sugiere Gould) de modo que no hay ningún tipo de ayuda que podamos encontrar entre los conceptos de la biología. Esto es seguramente lo que muchos humanistas y científicos sociales esperan ardientemente, pero también es altamente improbable, por las razones que ya hemos visto. Entre los dos extremos se encuentran probables y valiosas perspectivas: que haya una grande (o más grande) e importante (o simplemente medianamente interesante) transferencia de conceptos de la biología a las ciencias humanas. Podría ser, por ejemplo, que aunque los procesos de la transmisión cultural de ideas son verdaderos *fenómenos* darwinianos, se resistan por varias razones a ser atrapados en una *ciencia* darwiniana, de modo que tendríamos que optar por las realizaciones «meramente filosóficas» que pudiéramos espigar, dejando a la ciencia abordar otros proyectos.

Consideremos, en primer lugar, la afirmación de que los fenómenos de la evolución cultural son verdaderamente darwinianos. Entonces podemos

retornar a las complicaciones escépticas. Al comienzo, la perspectiva del meme es evidentemente perturbadora, incluso produce consternación. La podemos resumir con un eslogan: «Un erudito es justamente el camino a través del cual una biblioteca hace otra biblioteca». Yo no sé lo que pensarán ustedes, pero en principio no me siento atraído por la idea de que mi cerebro sea una especie de estercolero en el cual las larvas de las ideas de otras personas se renovaran, antes de enviar copias de sí mismas en una especie de diáspora informativa. Me parece algo así como despojar a mi mente de su importancia en tanto que autor y crítico. ¿Quién es el responsable, de acuerdo con este modo de ver las cosas, nosotros o nuestros memes?

No hay una respuesta simple para esta importante cuestión. No puede haberla. Nos gustaría vernos a nosotros mismos como creadores de ideas, parecidos a dioses, manipulándolas y controlándolas a nuestro antojo y juzgándolas desde una independiente y olímpica posición. Pero, aunque éste sea nuestro ideal, sabemos que no es la realidad, ni siquiera en las mentes más magistrales y creativas. Es bien conocido lo que Mozart pensaba de sus propias ideas:

> Cuando me siento bien y con buen humor, o cuando voy de excursión o paseo después de una buena comida, o por la noche cuando no puedo dormir, las ideas se aglomeran en mi mente fácilmente. ¿De dónde y cómo llegan? No lo sé y *no tengo nada que ver con esto*. Las que me gustan, las conservo en la memoria y las canturreo; al menos otros me han dicho que eso suelo hacer[122]. (La cursiva es mía).

Mozart se encuentra en buena compañía. Raro es el novelista *que no* afirma que los personajes han sido «sacados de su propia vida»; los pintores tienden todavía más a confesar que sus cuadros se independizan y se pintan a sí mismos y los poetas admiten humildemente que son los sirvientes o incluso los esclavos de las ideas que hierven en sus cabezas, no los jefes. Podemos citar casos de memes que persisten en nuestras mentes sin ser dirigidos ni apreciados, o que se extienden —como los rumores— a pesar de la general desaprobación de esta expansión por parte de aquellos que contribuyen a su propagación.

El otro día estaba desconcertado —consternado— al descubrirme tarareando una melodía mientras paseaba. No era un tema de Haydn, de Brahms ni de Charlie Parker, ni siquiera de Bob Dylan: estaba tarareando enérgicamente «It Takes Two to Tango» (un *bit* pegadizo para los oídos, perfectamente miserable y totalmente irrecuperable que fue inexplicablemente popular durante algún tiempo, en la década de los cincuenta). Estoy seguro de que nunca en mi vida habría escogido esta melodía, estimado esta melodía ni, de algún modo, juzgado que esta melodía era mejor que el silencio, pero allí

estaba ese horrible virus musical, tan robusto en mi *pool* de memes, al menos tanto como cualquier otra melodía que realmente estimo. Y ahora, para empeorar las cosas, he resucitado este virus en muchos de vosotros, por lo que no dudaríais en maldecirme si en el futuro me sorprendieran tarareando, por vez primera en treinta años, la aburrida tonada.

El lenguaje humano, primero hablado y después, muy recientemente, escrito es, con toda seguridad, el principal medio de transmisión cultural, pues crea la *infosfera* en la cual se produce la evolución cultural. Hablar y escuchar, escribir y leer, éstas son las subyacentes tecnologías de la transmisión y replicación que más se parecen a las tecnologías del ADN y del ARN en la biosfera. No necesito aburrirme revisando hechos bien conocidos en relación con la reciente proliferación explosiva de estos medios a través de los memes mediante tipos móviles de imprenta, radio y televisión, xerografía, ordenadores, faxes y correo electrónico. Todos somos conscientes de que hoy vivimos en un mar rebosante de memes transmitidos por papel, respirando en una atmósfera de memes transmitidos electrónicamente. Los memes se extienden alrededor del mundo a la velocidad de la luz y se replican a un ritmo que hace que incluso la mosca de la fruta y las células de la levadura parezcan, en comparación, congeladas. Saltan promiscuamente de vehículo en vehículo, de medio en medio, y demuestran que virtualmente es imposible someterlos a cuarentena.

Los genes son invisibles; conducidos por vehículos de genes (organismos) en los cuales tienden a producir efectos característicos (efectos fenotípicos) sus destinos están, a largo plazo, determinados. Los memes también son invisibles, y viajan en vehículos de memes: imágenes, libros, dichos (en particular lenguajes, orales o escritos, sobre papel o magnéticamente codificado, etc.). Herramientas y edificios y otras invenciones son también vehículos de memes (Campbell 1979). Un vagón con ruedas radiadas transporta no sólo grano o carga de un lugar a otro, sino que también transporta, de mente en mente, la brillante idea de un vagón con ruedas radiadas. La existencia de un meme depende de su incorporación a algún medio; si se destruyen estas incorporaciones físicas, este meme se extingue. Puede, naturalmente, hacer una subsiguiente e independiente reaparición, del mismo modo que podrían, en principio, los genes de los dinosaurios, aparecer juntos —genes y dinosaurios— en algún distante futuro, pero los dinosaurios que esos genes crearían y habitarían no serían los descendientes de los dinosaurios originales, de modo tan directo como lo somos nosotros. La suerte de los memes está determinada de modo similar por el hecho de que

sus copias y copias de copias puedan persistir y multiplicarse, y esto depende de las fuerzas selectivas que actúan directamente sobre los diversos vehículos físicos que los incorporan.

Los memes, como los genes, son potencialmente inmortales, pero, al igual que los genes, dependen de la existencia de una cadena continua de vehículos físicos que persiste frente a la segunda ley de la termodinámica. Los libros son relativamente permanentes y las inscripciones en los monumentos son aún más permanentes, pero a menos que se encuentren bajo la protección de conservadores humanos, tienden a disolverse con el tiempo. Manfred Eigen ha reflexionado sobre este mismo punto acerca de los genes, aunque llevando la analogía en otra dirección:

> Consideremos, por ejemplo, una de las composiciones de Mozart, una que se conserve de forma estable en nuestro repertorio de conciertos. La razón de que se conserve no es que las notas de esta obra estén impresas con una tinta especialmente duradera. La persistencia con la que una sinfonía de Mozart reaparece en nuestros programas de concierto es solamente una consecuencia de su elevado valor de selección. Para mantener su efecto, la obra debe ser interpretada una y otra vez, el público debe tomar nota de ello y debe ser continuamente evaluada en competición con otras composiciones. La estabilidad de la información genética tiene causas similares (Eigen 1992:15).

Como sucede con los genes, la inmortalidad es más un asunto de replicación que de longevidad de los vehículos individuales. Como vimos, la preservación de los memes platónicos, a través de una serie de copias de copias, es un caso especialmente impresionante. Mientras unos pocos fragmentos de papiros de los textos platónicos, casi contemporáneos de Platón, aún existen, la supervivencia de sus memes casi no debe nada a la estabilidad química de estos fragmentos. Hoy las bibliotecas contienen miles si no millones de copias físicas (y traducciones) de *La República* de Platón y los ancestros claves en la transmisión de este texto se convirtieron hace ya siglos en polvo.

La simple replicación física de los vehículos no es suficiente para asegurar la longevidad de los memes. Unos miles de copias de un nuevo libro, encuadernadas en pasta dura, pueden desaparecer sin dejar huella en pocos años, y ¿quién sabe cuantas brillantes cartas al editor, reproducidas en cientos de miles de copias, desaparecen cada día dentro de basureros e incineradoras? Puede llegar un día en el que evaluadores no humanos de los memes sean suficientes para seleccionar y ordenar la preservación de memes concretos pero, de momento, los memes siguen dependiendo, al menos indirectamente, de uno o más de sus vehículos pasando, al menos, una breve fase larvaria en un notable tipo de nido para memes: la mente humana.

Hay limitación de mentes y cada mente tiene una capacidad limitada para

los memes por lo cual existe una considerable competición entre los memes por entrar en el mayor número de mentes posible. Esta competición es la mayor fuerza de selección en la infoesfera, y, del mismo modo que en la biosfera, los retos se han enfrentado con gran ingenio. ¿Ingenio de quién?, nos gustaría preguntar, pero por ahora debemos saber que ésta no es siempre una buena pregunta; la ingeniosidad está *allí* para ser apreciada, cualquiera que sea la fuente. Como un virus insensato, las perspectivas de un meme dependen de su diseño, no de su diseño «interno», sea cual sea, sino del diseño que muestra al mundo, de su fenotipo, que es la vía por la que afecta a las cosas de su entorno. Las cosas en su entorno son mentes y otros memes.

Por ejemplo, cualesquiera que sean las virtudes (desde nuestra perspectiva) que tengan los siguientes memes, todos tienen en común la propiedad de poseer expresiones fenotípicas que tienden a hacer su propia replicación más verosímil, mediante la inutilización o el prevaciamiento de las fuerzas ambientales que tenderían a extinguirlos: así, por ejemplo, el meme de la *fe* desalienta el ejercicio del juicio crítico capaz de decidir que la idea de la fe es una idea peligrosa (Dawkins 1976:212); los memes de la *tolerancia* o de la *libre expresión*; el meme para incluir *un aviso en una cadena de cartas* acerca de las terribles consecuencias que sufrieron aquellos que rompieron la cadena en el pasado; el meme de la *teoría de la conspiración*, construido en respuesta a la objeción de que no hay buena evidencia de conspiración: «¡Desde luego que no, esto demuestra lo poderosa que es la conspiración!». Algunos de estos memes son «buenos» y otros «malos»; lo que tienen en común es el efecto fenotípico que sistemáticamente tiende a inutilizar las fuerzas selectivas desplegadas contra ellos. En igualdad de condiciones, la memética predice que los memes de la teoría de la conspiración persistirán, independientemente de su verdad, y que el meme de la fe está capacitado para asegurar su propia supervivencia, y la de los memes religiosos que arrastra, incluso en los ambientes más racionalistas. Naturalmente el meme de la fe exhibe una *aptitud* para desenvolverse en el medio ambiente que es *frecuencia-dependiente*: florece especialmente en compañía de memes racionalistas. En un mundo poco escéptico, el meme de la fe no atrae bastante atención y, en consecuencia, tiende a estar durmiente en las mentes y rara vez se introduce en la infoesfera. (¿Podemos demostrar entre los memes de la fe y los memes de la razón los clásicos ciclos *boom*-colapso en la población predador-presa? Probablemente no, pero puede ser instructivo examinar esta cuestión y preguntarnos por qué no).

Otros conceptos de la genética de la población se transfieren con facilidad. Aquí encontramos un caso de lo que un genetista llamaría un *linked loci* (‘dos locus ligados’): dos memes que resultan estar físicamente ligados de tal modo que siempre tienden a replicarse juntos, una tendencia que afecta a sus oportunidades. Hay una magnífica marcha ceremonial, muy conocida por muchos de nosotros y generalmente admirada. Es excitante, brillante y grandiosa, precisamente el tipo de música, podríamos pensar, digna de ser utilizada en las inauguraciones, bodas y otras ocasiones festivas, quizá llevando a «Pompas y Circunstancias» y a la Marcha Nupcial de Lohengrin casi a la extinción, si no fuera por el hecho de que su meme musical está demasiada ligado a su título, en el que pensaríamos tan pronto escucháramos la música: la no utilizable obra maestra de sir Arthur Sullivan, «Presentando al señor verdugo mayor del reino». Si esta marcha no tuviera letra y se titulara, por decir algo, «La marcha de Koko», su interpretación no hubiera sido descalificada. Pero el título real, que corresponde a las cinco primeras palabras de su letra, las cuales están fuertemente ligadas con la melodía, garantizan virtualmente una *cadena de pensamientos* en la mayoría de los oyentes que resultaría indeseable en cualquier ocasión festiva. Este es el efecto fenotípico que previene una mayor replicación de este meme. Si las interpretaciones de *El Mikado* fueron declinando a lo largo de los años, hasta que llegó un tiempo en el que poca gente conocía la letra de la marcha, para no hablar de la estúpida historia, la marcha podría volver por sí misma como una pieza de música ceremonial sin palabras, ¡excepto por su maldito título! ¿No creen que parecería buena si se ubicara en el programa precisamente antes del discurso del vicecanciller a los graduados?

Este es realmente un caso llamativo de uno de los fenómenos más importantes de la infoesfera: la colocación de un filtro equivocado que conduce a la eliminación de memes debido a las relaciones que se establecen a partir de ellos. Existe incluso un meme que da nombre a este fenómeno: *arrojar al niño junto con el agua de la tina*. Este libro tiene como intención prioritaria deshacer los desafortunados efectos provocados por someter a un filtrado erróneo a los memes darwinianos, un proceso que ha venido teniendo lugar incluso desde que el propio Darwin se confundió acerca de cuáles eran sus mejores ideas (aunque alguno de sus enemigos estuvieron de acuerdo con ellas) y cuáles eran las peores (aunque aquellas parecían realizar una especie de servicio de guardia contra ciertas perniciosas doctrinas). (R. Richards 1987 presentó una historia especialmente fascinante de la evolución de las ideas de la evolución). Todos tenemos filtros de esta clase: «Yo ignoro todo lo que

aparece en X». Para algunos X es el *National Geographic* o *Pravda*; para otros es *The New York Review of Books*; todos jugamos con la suerte, contando con que las «buenas» ideas pueden llegar al primer plano de nuestra atención a través de la serie de filtros colocados por otros.

La estructura de los filtros es, en sí misma, un meme de considerable solidez. John McCarthy, uno de los fundadores de la inteligencia artificial (y acuñador de su nombre, un meme en sí mismo, base independiente en la infoesfera) sugirió, en cierta ocasión, ante una audiencia humanística, que la red de correo electrónico revolucionaría la ecología del poeta. Tan sólo unos cuantos poetas pueden ganarse la vida vendiendo sus poemas —decía McCarthy— debido a que los libros de poemas son volúmenes delgados y caros, comprados por un número escaso de gente y de bibliotecas. Pero imaginemos lo que sucedería si los poetas presentaran sus poemas en una red internacional, donde cualquiera pudiera leerlos o copiarlos por un penique transferido electrónicamente a la cuenta corriente abierta para los derechos de autor del poeta. Esta sería una fuente segura de ingreso para muchos poetas, sostenía McCarthy. Independientemente de cualquier objeción estética a que los poetas y los amantes de la poesía pudieran incorporar sus poemas en los medios de comunicación electrónicos, la contra hipótesis obvia surge de la memética de la población. Si se estableciera esta red electrónica, ningún amante de la poesía querría leer tediosamente miles de archivos electrónicos repletos de poemas triviales que quieren pasar por buenos; sería un nicho creado por varios memes actuando como filtros de poesía. Uno se puede subscribir, por unos cuantos peniques, a un servicio editorial que haga un recorrido electrónico por la infoesfera a la búsqueda de buenos poemas. Florecerían así diferentes servicios con diferentes criterios, como servicios para revisar los diferentes servicios y servicios que revisarían, coleccionarían, darían formato y presentarían las obras de los mejores poetas en delgados volúmenes electrónicos, de los cuales sólo se comprarían unos pocos. En otras palabras, los memes para la edición y la crítica encontrarían nichos en cualquier ambiente de la infoesfera; florecerían debido a la corta reserva y limitada capacidad de la mente, cualquiera que sea el medio de transmisión entre las mentes. ¿Dudamos de esta predicción? Si es así, me gustaría discutir la construcción de una apuesta apropiada. Aquí, de nuevo, como hemos visto en el pensamiento evolucionista la explicación procede de la hipótesis de que los procesos —cualquiera que sea el medio y cualquiera que sea el zigzag contingente de sus particulares trayectorias— se dirigirán hacia los movimientos forzados y otros buenas mañas en el espacio pertinente.

La estructura de los filtros es compleja y de respuesta rápida frente a los nuevos retos, pero es natural que no siempre funcione. La competición entre los memes para abrir una brecha en los filtros conduce a una «carrera de armamentos», de maniobras y contramaniobras, incluso con una más elaborada «propaganda» desarrollada contra más capas de filtros selectivos. En la dignificada ecología académica, no llamamos a esto propaganda, pero la misma carrera armamentística se manifiesta en los membretes de los departamentos, «referencias ciegas», proliferación de revistas especializadas, revisiones de libros, revisiones de revisiones de libros y antologías de «obras clásicas». Estos filtros no siempre intentan preservar lo mejor. Los filósofos, por ejemplo, pueden preguntarse cuántas veces son cómplices del aumento de audiencia para un artículo de segunda categoría, simplemente porque su curso de introducción necesita una versión sincera de una mala idea, que incluso un novato pueda refutar. Algunos de los artículos reproducidos con más frecuencia en la filosofía del siglo XX son famosos precisamente debido a que nadie cree en ellos; cualquiera puede ver dónde está el error[123].

Un fenómeno relacionado con la competición de los memes que tratan de captar nuestra atención es el *feedback* positivo. En biología este fenómeno se manifiesta en la «selección sexual abrumadora» que explica la larga y molesta cola del ave del paraíso o del pavo real (para detalles, véanse Dawkins 1986a:195-220, Cronin 1991 y Matt Ridley 1993). Dawkins aporta un ejemplo extraído del mundo editorial: «Las listas de libros más vendidos se publican semanalmente y es sin duda verdad que tan pronto como un libro vende bastante ejemplares y aparece en una de estas listas, sus ventas aumentan aún más, simplemente en virtud de este hecho. Los editores hablan del “despegue” de un libro y aquellos editores que tienen algún conocimiento de la ciencia incluso hablan de “masa crítica para el despegue”» (1986a:219).

Los vehículos de los memes habitan nuestro mundo junto con toda su fauna y flora, grande y pequeña. Sin embargo, de manera general, solamente son «visibles» para la especie humana. Consideremos el medio ambiente que rodea al tipo habitual de paloma de Nueva York, cuyos ojos y oídos se ven asaltados cada día por, aproximadamente, las mismas palabras, imágenes y otros signos y símbolos que asaltan a cada neoyorquino. Estos vehículos físicos de memes pueden interferir de modo importante en el bienestar de las palomas, pero no debido a los memes que transportan; no significa nada para la paloma que la migaja que ha encontrado esté debajo de una página de *The National Enquirer* o del *The New York Times*. Para los seres humanos, por otro lado, cada vehículo de memes es un amigo o enemigo en potencia, con el

beneficio de un regalo que incrementará nuestros poderes o como un caballo regalado que nos distraerá recargando nuestra memoria y alterando nuestro juicio.

3. *¿Podría existir una ciencia de la memética?*

> La magnitud de la tarea me impresiona hasta dejarme perplejo. Pero, sobre todo, si se acepta la perspectiva evolucionista, los intentos de discutir la ciencia (u otro tipo de actividad conceptual) se hacen bastante más difíciles, tan difíciles como para provocar una parálisis total.
>
> David Hull, «*The Naked Mente*»

> Los memes son capaces de dar instrucciones, no para sintetizar proteínas como hacen los genes, sino dirigidas a comportamientos. Sin embargo, los genes pueden hacer esto muy indirectamente a través de la síntesis de proteínas. Por otro lado, la replicación de los memes, al implicar modificaciones neuroestructurales, se asocia, invariablemente, con la inducción de síntesis proteica.
>
> Juan Delius, «*The Nature of Culture*»

Todo esto es muy estimulante, pero estamos encubriendo un buen número de complicaciones. Puedo escuchar un coro de escépticos tratando de construir su oportunidad. ¿Recordáis la historia recogida al final del capítulo 4, en la que relataba la displicente opinión de Francis Crick sobre la genérica de poblaciones como ciencia? Si la genética de las poblaciones es calificada por los pelos como ciencia —y ciencia obsoleta en esa historia—, ¿qué probabilidades existen para una verdadera *ciencia* de la memética? Los filósofos, dirá alguno, pueden apreciar esta (aparente) intuición como el hallazgo de una nueva perspectiva, pero si no podemos convertirla en ciencia real, con hipótesis que se puedan someter a prueba, formalizaciones fiables y resultados cuantitativos, en realidad, ¿para qué sirve todo esto? El mismo Dawkins nunca ha dicho que hubiera fundado con la memética una nueva disciplina científica. Si esto es así, ¿hay algo erróneo en el concepto de meme?

¿Lo que mantiene a un meme como ADN, mantiene a un gen? Varios comentaristas (véase, por ejemplo, Delius 1991) han argumentado a favor de la identificación de los memes con estructuras cerebrales complejas, en paralelismo con la identificación de los genes con complejas estructuras de ADN. Pero, como ya hemos visto, es un error *identificar* a los genes con sus vehículos en el ADN. La idea de que la evolución es un proceso algorítmico

es la idea de que para este proceso debe haber una útil descripción en términos neutrales respecto al sustrato. Como propuso George Williams hace muchos años: «En la teoría de la evolución, un gen puede ser definido como cualquier *información* hereditaria para la cual existe un sesgo hacia una selección favorable o desfavorable que es igual a varias o muchas veces su ritmo endógeno de cambio» (1966:25; la cursiva es mía). La importancia de la separación entre información y vehículo es todavía más fácil de comprender en el caso de los memes[124]. El obvio problema señalado por todos es que es muy improbable —pero no completamente imposible— que haya un «lenguaje cerebral» uniforme en el cual se almacene la información en los diferentes cerebros humanos, lo que hace a los cerebros muy diferentes de los cromosomas. Los genetistas han identificado recientemente una estructura cromosómica que denominan el *homeobox*; a pesar de las diferencias, esta estructura es identificable en especies animales muy alejadas —quizás en todas— de modo que es muy antigua, y tiene un papel central en el desarrollo embriológico. Puede sorprendernos, en principio, que un gen identificado en el homeobox del ratón por cumplir un papel muy importante en el desarrollo del ojo, ha deletreado casi siempre el mismo codón que un gen apellidado *sinojos* por su efecto fenotípico, que ha sido identificado en el homeobox de la mosca *Drosophila* de la fruta. Pero más nos asombraría descubrir que el complejo celular cerebral que había almacenado el meme de Benjamin Franklin correspondiente a las gafas bifocales (que él inventó) era el mismo o muy similar al complejo celular cerebral que se utiliza hoy para almacenar el meme de las bifocales dondequiera que cualquier niño en Asia, África o Europa tiene noticias de las gafas bifocales por primera vez, leyendo algún texto acerca de ellas, viéndolas en televisión o bien observándolas colocadas sobre la nariz de sus padres. Lo que esta reflexión deja claro es el hecho de que lo que se preserva y transmite en la evolución cultural es *información*, en el sentido de un medio y un lenguaje neutrales. De este modo, el meme es primariamente una clasificación *semántica* y no una clasificación *sintáctica* que pueda directamente observarse en el «lenguaje cerebral» o lenguaje natural.

En el caso de los genes tenemos la suerte de encontrarnos con un gratificante y fuerte alineamiento de la identidad semántica con la sintáctica: *hay* un único lenguaje genético, en el cual el significado se preserva (aproximadamente) a lo largo de todas las especies. Con todo, es importante distinguir los tipos semánticos de los tipos sintácticos. En la Biblioteca de Babel identificamos un conjunto de variantes sintácticas de textos dentro de la

galaxia *Moby Dick*, en razón de lo que nos dicen *acerca de algo*, no de su similitud sintáctica. (Pienso en las diferentes traducciones de *Moby Dick* y también en versiones reducidas en inglés, esquemas y ayudas para su estudio, por no decir nada de las versiones cinematográficas y en otros medios de información). Nuestro interés por identificar y reidentificar genes en las edades de la evolución es del mismo modo *primario* debido a la uniformidad de los efectos fenotípicos, lo que ellos son *acerca de algo* (sea la construcción de la hemoglobina o de los ojos). Nuestra habilidad para confiar en su identificabilidad sintáctica en el ADN es un avance reciente, e incluso cuando no cabe concebiblemente aprovecharnos de ello (deduciendo, por ejemplo, hechos en relación con cambios genéticos a partir de lo que podemos observar en el registro fósil de las especies que no han dejado ADN para que nosotros lo podamos «leer») podemos aún confiadamente hablar de genes, es decir, de la información que debe haber sido preservada o Transmitida.

Es concebible, pero difícilmente probable y ciertamente no necesario, que algún día se descubra una identidad impresionante entre las estructuras cerebrales que almacenan la misma información, permitiéndonos identificar sintácticamente los memes. Sin embargo, aunque encontráramos tal improbable bendición del cielo, debemos asirnos al más abstracto y fundamental concepto de los memes, dado que ya sabemos que la transmisión y almacenamiento de los memes puede realizarse indefinidamente en estructuras no cerebrales —en artefactos de todo tipo— que no dependen de un lenguaje descriptivo compartido. Si hubiera, incluso, una transmisión «multimedia» y una transformación de la información, se trataría de información y transmisión cultural. De este modo, algunas de las variantes del triunfo reduccionista que cabe esperar en la biología —por ejemplo, el descubrimiento de cómo diferentes formas de la hemoglobina son «deletreadas» en todas las especies en el mundo— son descartadas, casi con certeza, de cualquier ciencia de la cultura, a pesar de las profecías sobre una edad de oro de la lectura de la mente que suelen hacer en estos días los ideólogos de la neurociencias.

Esto frustraría tan sólo a algunas clases de ciencia memética, pero ¿no es realmente la situación todavía peor? Como hemos visto, la evolución darwiniana depende de copias de *muy* alta fidelidad, unas copias casi perfectas, gracias a la exquisita maquinaria para la lectura de pruebas y la duplicación de los lectores del ADN que acompaña a los textos de ADN. Si se eleva mínimamente la tasa de mutación, la evolución se desquicia; la selección natural no puede trabajar para garantizar a largo plazo la

adaptabilidad al medio ambiente. Por otro lado, las mentes (o cerebros) no son, en absoluto, algo parecido a fotocopiadoras. Por el contrario, en lugar de hacer pasar sus mensajes exactamente, corrigiendo tan sólo los errores tipográficos, cuando existen, el cerebro parece estar diseñado para hacer precisamente lo contrario: transformar, inventar, interpolar, censurar y, generalmente, hacer mezclas con el *input* antes de producir algún *output*. ¿No es precisamente uno de los puntos clave de la evolución y de la transmisión cultural el extraordinario y elevado ritmo de la mutación y la recombinación a nivel cerebral? Según parece, *raras veces* hacemos pasar un meme sin que sea modificado, a menos que se trate de aprendices rutinarios con mentes propicias a la literalidad. (¿Acaso somos fanáticas enciclopedias ambulantes?). Además, como ha subrayado Steven Pinker (comunicación personal), bastantes mutaciones de las que suceden en los memes —en qué cuantía no está claro— son manifiestamente mutaciones *dirigidas*: «Los memes, como la teoría de la relatividad, no son el producto acumulativo de millones de mutaciones *aleatorias* (no dirigidas) de alguna idea original, sino que en la cadena de producción cada cerebro añade al producto, de un modo no aleatorio, una enorme masa de valor». Naturalmente, el poder total de la mente cuando un meme anida proviene de lo que un biólogo llamaría *cruce de linajes* o *anastomosis* (la recuperación conjunta de pools de genes separados). Como ha señalado Gould, «Las topologías básicas de los cambios biológicos y culturales son completamente diferentes. La evolución biológica es un sistema en constante divergencia, sin una posterior unión de las ramas. Los linajes, una vez diferenciados, se separan para siempre. En la historia humana, por el contrario, la transmisión entre linajes es, quizá, la mayor fuente de cambio cultural» (1991a:65).

Por otra parte, cuando los memes entran en contacto unos con otros en una mente, poseen una maravillosa capacidad para acoplarse rápidamente, modificando sus efectos fenotípicos para adaptarse a las circunstancias; entonces la receta del nuevo fenotipo consigue replicarse y la mente emite o publica los resultados de la mezcla. Por ejemplo, mi nieto de tres años, a quien le encanta la maquinaria de construcción, nos sorprendió no hace mucho tiempo con una interesante mutación de una canción de cuna: «¡Pum!, se va el diesel». «¡Pop!, goes the weasel» (*¡Pum!, se va la comadreja*) es el estribillo de una famosa canción de cuna. Mi nieto no se dio cuenta de lo que había hecho (cambiar *weasel* por *diesel*) pero yo, a quien nunca se le había ocurrido la frase, compruebo ahora que este meme de mi nieto se ha replicado. Como en el caso de los chistes, discutido previamente, este

modesto momento de creatividad es una mezcla de chiripa y apreciación, distribuida en varias mentes, ninguna de las cuales posee la autoría de una creación especial. Es algo así como una replicación lamarckiana de características adquiridas, como Gould y otros han sugerido[125]. La gran creatividad y actividad de las mentes humanas como domicilios temporales de los memes parece garantizar que las líneas de descenso estén inevitablemente enredadas y que los fenotipos (los «diseños corporales» de los memes) cambien con tanta rapidez que exista la posibilidad de que se mantengan los llamados «tipos naturales». Recuperemos del capítulo 10 la idea de que las especies son invisibles si no existe un mínimo de «estasis», pero recordemos, también, que esta es una cuestión metodológica y no metafísica: si las especies no fueran hasta cierto punto estáticas, no podríamos *investigar* y organizar los hechos que son necesarios para hacer ciertos tipos de ciencia; sin embargo, esto no demostraría que los fenómenos no fueran gobernados por la selección natural. De modo similar, la conclusión aquí sería una conclusión *epistemológicamente* pesimista: aun en el caso de que los memes se originasen por un proceso de «descenso con modificación» nuestras probabilidades de poner en marcha una ciencia que explorara el trayecto seguido por este descenso serían escasas.

Una vez que hemos mostrado nuestra preocupación de esta forma, nos queda por señalar lo que parece ser una solución parcial. Uno de los hechos más llamativos de la evolución cultural es la facilidad, fiabilidad y confianza con que podemos identificar cosas comunes a pesar de las enormes diferencias en los medios de comunicación a través de los cuales se difunden. ¿Qué tienen en común *Romeo y Julieta* y la película *West Side Story*? (Dennett 1987b). No una larga serie de personajes que se expresan en inglés, ni una secuencia de proposiciones (en inglés o, en las versiones dobladas, en francés, alemán, etc.). Lo que tienen en común, naturalmente, no es una propiedad *sintáctica*, ni un sistema de propiedades, sino una propiedad *semántica*, o sistema de propiedades: la historia, no el texto; los personajes y sus personalidades, no sus nombres y discursos. Lo que tan fácilmente identificamos como la misma cosa en ambos casos es la difícil situación en la que ambos autores, William Shakespeare y Arthur Laurents (que escribió el libro de *West Side Story*) desean que creamos. De este modo, es tan sólo al nivel de *objetos intencionales*, una vez que hemos adoptado la posición intencional, desde el que podemos describir estas propiedades comunes[126]. Cuando adoptamos esta posición, la búsqueda de características comunes a menudo se soporta como una piedra en un zapato.

¿Funciona esta ayuda? Sí, aunque debemos tener cuidado respecto a un problema que ya hemos identificado, con aspectos diferentes: el problema de definir lo que es un plagio (o préstamo respetuoso) de la evolución convergente. Como señala Hull (1982:300) no debemos considerar dos objetos culturales *idénticos* como ejemplos del mismo meme a menos que estén relacionados por la vía del descenso. (Los genes de los ojos del pulpo no son los mismos que los de los ojos del delfín, aunque los ojos parezcan similares). Esto puede crear un sinnúmero de ilusiones o de indeterminaciones en los evolucionistas culturales, cada vez que intenten seguir las huellas de los memes responsables de las buenas mañas. Mientras más abstracto es el nivel en el cual identificamos los memes, más difícil será definir la evolución convergente del descenso. Lo que pasa es que sabemos, porque ellos lo han dicho, que los creadores de *West Side Story* (Arthur Laurents, Jerome Robbins y Leonard Bernstein) extrajeron la idea de su obra de *Romeo y Julieta*, pero si hubiesen guardado el secreto, podíamos haber pensado que simplemente habían reinventado la rueda, redescubierto un «universal» de la cultura, que aparecería, por sí mismo, en casi toda evolución cultural. Mientras más puramente semánticos sean nuestros principios de identificación —o, en otras palabras, lo menos ligados a formas particulares de expresión— más difícil será seguir con confianza las huellas del descenso. (Recordemos (capítulo 6) que fueron peculiaridades en la forma especial de expresión las que permitieron a Otto Neugebauer encontrar la clave crucial para descifrar el misterio de la traducción de la efemérides de Babilonia). En la ciencia de la cultura, éste es el mismo problema epistemológico con que el que se enfrentan los taxonomistas, cuando, en el análisis cladista (Mark Ridley 1985), tratan de separar homología de analogía, caracteres ancestrales de caracteres derivados. Idealmente, en el terreno imaginado por los cladistas culturales, uno desearía encontrar «caracteres» —literalmente, caracteres alfabéticos— que fueran funcionalmente elecciones opcionales dentro de un enorme número de posibles alternativas. Si encontramos la totalidad de los parlamentos de Antonio y María que, sospechosamente, replican las palabras y las frases de Romeo y de Julieta, no necesitaríamos las claves autobiográficas de Laurents, Robbins o Bernstein. No dudaríamos en declarar que la coincidencia de las palabras no era coincidencia; el espacio de diseño es demasiado vasto para que esto sea creíble.

Sin embargo, en nuestros intentos hacia una ciencia de la evolución cultural, no podemos contar, generalmente, con tales descubrimientos. Supongamos, por ejemplo, que deseamos argumentar que instituciones tales

como la agricultura o la monarquía, o incluso prácticas concretas como el tatuaje o el estrechar las manos, descienden de un ancestro cultural común, en lugar de haber sido reinventadas independientemente. Hay un intercambio buscando el mayor beneficio. En tanto que hemos de llegar hasta niveles funcionales abstractos (o semánticos) para encontrar nuestras características comunes, perdemos la capacidad de distinguir homología de analogía, descenso de evolución convergente. Esta situación ha sido tácitamente apreciada por los estudiosos de la cultura, de modo independiente del pensamiento darwiniano. Consideremos, por ejemplo, lo que podemos deducir de unos restos de cerámica. Los antropólogos que buscan evidencias de culturas compartidas quedan más impresionados por las comunes idiosincrasias del estilo decorativo que por las formas funcionalmente comunes. O bien consideremos el hecho de que dos culturas muy lejanas utilizaron barcos, este dato no evidencia, en modo alguno, una herencia cultural compartida. Si ambas culturas hubieran pintado ojos en la proa de sus barcos, esto ya sería más interesante, aunque aún seguiría siendo un movimiento obvio en el juego del diseño. Si ambas culturas hubieran pintado en las proas de sus barcos, por decir algo, hexágonos azules, este hecho nos estaría diciendo algo, naturalmente.

El antropólogo Dan Sperber, que ha reflexionado ampliamente sobre la evolución cultural, piensa que hay un problema cuando se plantea la utilización de objetos intencionales abstractos como anclajes de un proyecto científico. Tales objetos abstractos, dice Sperber,

no entran directamente en relaciones causales. Lo que causó nuestra indigestión no fue la receta de la salsa Mornay en abstracto, sino nuestro anfitrión que tras leer una descripción pública, de la salsa, se había formado una representación mental y la había seguido más o menos con éxito. Lo que causó el divertido miedo del niño no fue el cuento de Caperucita Roja en abstracto sino lo que entendió de las palabras de su madre. Insistiendo en la cuestión, lo que causó que la receta de la salsa Mornay y la historia de Caperucita Roja llegaran a ser representaciones culturales no son, o al menos no lo son directamente, sus propiedades formales, sino la construcción de millones de representaciones mentales casualmente ligadas por millones de descripciones públicas (Sperber 1985:77-78).

Lo que Sperber dice acerca del carácter indirecto del papel de las características abstractas es verdad, pero lejos de ser un obstáculo para la ciencia, es la mejor clase de invitación para la ciencia: una invitación a cortar el nudo gordiano de la enmarañada causalidad con una formulación abstracta que es previsible precisamente *porque* ignora rodas aquellas complicaciones. Por ejemplo, los genes son seleccionados debido a sus efectos fenotípicos sólo visibles indirecta y estadísticamente. Consideremos la siguiente predicción: dondequiera que encontremos polillas con camuflajes en las alas,

hallaremos que alrededor de ellas existen predadores de visión aguda y dondequiera que encontremos polillas que resultan masivamente cazadas por murciélagos con sistema de localización mediante eco, encontraremos que las polillas han desarrollado en sus alas un camuflaje mediante un dispositivo de interferencia de los ecos o un talento especial para crear figuras de vuelo que les permitan huir. Naturalmente, nuestro último objetivo es explicar en profundidad cualquier característica que encontremos en las polillas y sus alrededores, hasta llegar al mecanismo molecular o atómico responsable, pero no hay razón para pedir que tal reducción sea uniforme o generalizable. La gloria de la ciencia radica en que puede encontrar los modelos a pesar del ruido (Dennett 1991b).

Las peculiaridades de la psicología humana (y de la digestión humana, como en el ejemplo de la salsa Mornay) son *eventualmente* importantes, pero no impiden un análisis científico de los fenómenos tratados. En efecto, como el mismo Sperber ha argumentado persuasivamente, podemos utilizar principios del más alto nivel como palancas para mantener abiertos los secretos a niveles más bajos. Sperber subraya la importancia de la invención de la escritura, que inició las grandes cambios en la evolución cultural. Demostró cómo se ha de razonar desde los hechos de la cultura preliteraria a los hechos de la psicología humana. (Sperber prefiere pensar que la evolución cultural se transmite a lo largo de líneas de *epidemiología* más que siguiendo líneas de *genética*, aunque la dirección de su teoría es casi la misma que la de Dawkins, hasta el punto de ser casi indistinguibles cuando se piensa acerca de lo que aparenta el tratamiento darwiniano de la epidemiología; véase Williams y Nesse 1991). He aquí la «ley de la epidemiología de las representaciones» de Sperber:

> En una tradición oral, todas las representaciones culturales son fácilmente recordadas; las representaciones difíciles de recordar se olvidan, o se transforman en otras más fácilmente recordables, antes de alcanzar un nivel cultural de la distribución (Sperber 1985:86).

Al principio parece trivial, pero veamos cómo podemos aplicarla. Cabe utilizar la existencia de un tipo particular de representación cultural endémica a la tradición oral para aclarar cómo trabaja la memoria humana, preguntando qué hay en *este* tipo de representación que la hace más memorable que otras.

Sperber señala que la gente es mejor recordando una anécdota que recordando un texto, al menos hasta hoy, ya que ahora la tradición oral se está desvaneciendo[127]. Pero aun hoy, a veces recordamos involuntariamente el estribillo de un anuncio, incluyendo sus precisas propiedades rítmicas, el

«tono de voz» y muchas otras características de «bajo nivel». Cuando los científicos deciden sobre un acrónimo o sobre eslóganes atractivos para denominar sus teorías, esperan, de este modo, hacerlas más memorables, convertirlas en memes más brillantes y atractivos. Y como consecuencia, los detalles reales de la representación son, a veces, mejores candidatos para llegar a la calidad de meme que el propio contenido representado. El uso de acrónimos es, en sí mismo, un meme —un metameme, naturalmente—, lo que se comprende debido a su demostrado poder en la promoción de memes de contenido que han ayudado a diseñar los memes de nombres. ¿Qué podemos decir de los acrónimos, las rimas o los enérgicos eslóganes que hace que se comporten tan bien en las competiciones por penetrar agresivamente en la mente humana?

Este tipo de cuestiones saca provecho, como hemos visto muchas veces, de una estrategia fundamental, tanto en la teoría de la evolución como en la ciencia cognitiva. Donde la teoría de la evolución considera a la información como transmitida a través de canales genéticos, cualesquiera que sean, la ciencia cognitiva considera a la información transmitida a través de canales del sistema nervioso, cualesquiera que sean, además de los medios adyacentes, tales como el aire traslucido que tan bien transmite el sonido y la luz. Podemos esquivar nuestra ignorancia de los cruentos detalles mecánicos sobre cómo la información llega de A a B, al menos temporalmente, y concentrarnos en las implicaciones del hecho de que alguna información llegó allí y otra no llegó.

Supongamos que nos asignan la tarea de cazar a un espía, o a toda una trama de espías en el Pentágono. Supongamos que se sabía que la información sobre los submarinos nucleares estaba cayendo en manos del enemigo. Un método para cazar al espía sería insertar varios *bits* de información falsa, aunque creíble, en varios lugares dentro del Pentágono y observar cómo sale a la superficie, en qué orden, en Ginebra, Beirut o cualquier lugar donde se encuentre el mercado de los secretos. Variando las condiciones y circunstancias podemos construir gradualmente un elaborado diagrama de la ruta —las diversas estaciones en el camino y las transferencias y lugares de encuentro— hasta llegar al punto de arrestar, debidamente, a la trama de espías, aunque sin saber todavía cuál ha sido el medio de comunicación utilizado. ¿Fue la radio? ¿Fueron minúsculas piezas de película fotográfica, con una elevadísima *ratio* de reducción, adheridas a los documentos? ¿Señales con el código de banderas? ¿Acaso el agente memorizó los planos y

sencillamente atravesó desnudo la frontera, o bien tenía una descripción verbal en código Morse, oculta en el disco flexible de su ordenador?

Al finalizar, deseamos conocer las respuestas a todas esas preguntas pero, mientras tanto, hay muchas cosas que podemos hacer en el dominio del sustrato neutral de la pura transferencia de información. En la ciencia cognitive, por ejemplo, el lingüista Ray Jackendoff [1987, 1993] demostró el sorprendente poder de este método en sus ingeniosas deducciones acerca del número de niveles de representación, y sus poderes, que *deben* utilizarse en tareas tales como obtener información tanto de la luz que impacta nuestros ojos por todo el camino, como de lugares donde podemos hablar acerca de lo que vemos. Jackendoff no tiene que conocer los detalles de la neurofisiología (aunque está interesado en ella, al contrario de muchos otros lingüistas) para sacar conclusiones fiables acerca de la estructura de los procesos y de las representaciones que transforman.

Lo que aprendemos a este nivel abstracto es, por derecho propio, científicamente importante. Es, naturalmente, la base de todo lo que es importante. Nadie lo ha expuesto mejor que el físico Richard Feynman:

> ¿A nadie le inspira nuestro actual cuadro del universo? Este valor de la ciencia sigue sin ser cantado por cantantes: estamos limitados a oír no una canción o un poema, sino una conferencia vespertina sobre esta cuestión. Esta no es todavía una edad científica.

Quizás una de las razones de este silencio es que hemos de saber cómo leer la música. Por ejemplo, el artículo científico puede decir: «El contenido en fósforo radioactivo del cerebro de la rata disminuye a la mitad en un período de dos semanas». Bien, ¿qué significa esto?

Significa que el fósforo que había en el cerebro de la rata —y también en el mío y en el de todos nosotros— no es el mismo fósforo que se encontraba hace dos semanas. Significa que los átomos que se encontraban en el cerebro han sido reemplazados: algunos de los que allí estaban antes se han ido.

Entonces, ¿qué es nuestra mente? ¿Qué tienen que ver estos átomos con la consciencia? ¡Las patatas de la semana pasada! Ellas pueden *recordar* ahora lo que sucedía en mi mente hace un año; una mente que hace ya tiempo ha sido reemplazada.

Nótese que esa cosa que yo llamo mi individualidad es sólo un modelo o una danza; esto es lo que significa cuando uno descubre el tiempo que tardan los átomos del cerebro en ser reemplazados por otros. Los átomos llegan a mi cerebro, bailan una danza y se van; hay siempre nuevos átomos, pero siempre

ejecutando la misma danza, recordando cuál fue la danza de ayer (Feynman 1988:244).

4. *La importancia filosófica de los memes*

> Si somos quisquillosos y puristas en el uso de las palabras, la «evolución» cultural no es evolución en modo alguno, aunque puede haber lo suficiente en común entre ambas como para justificar alguna comparación de sus principios.
>
> Richard Dawkins, *El relojero ciego*

> Hay tanta razón para esperar que una práctica cultural transmitida entre feligreses incremente la aptitud de estos feligreses, como la que hay para esperar que un virus de la gripe transmitido de la misma forma, incremente esta aptitud.
>
> George Williams,
> *Natural Selection: Domains, Levels, and Challenges*

Cuando Dawkins introdujo los memes en 1976, describió su innovación como una extensión literal de la clásica teoría darwiniana. Desde entonces, Dawkins ha adoptado una actitud ligeramente defensiva. En su libro *El relojero ciego* (1986) hablaba de una analogía «que yo encuentro inspiradora aunque puede ir a parar demasiado lejos si no somos cuidadosos». ¿Por qué se bate en retirada? ¿Por qué el meme del meme ha sido tan poco discutido dieciocho años después de la aparición del libro *El gen egoísta*?

En *The Extended Phenotype* (1982), Dawkins replica vigorosamente a la tormenta de críticas de los sociobiólogos y de otros, mientras admite algunos interesantes desajustes en las analogías entre genes y memes:

> Los memes no están ampliamente separados a lo largo de cromosomas lineales, y no está claro que ocupen y compitan por lugares concretos o que tengan alelos identificables… El proceso de copia es probablemente bastante menos preciso que el de los genes… Los memes pueden mezclarse parcialmente unos con otros de un modo que los genes no hacen (*ibidem*: 112).

Pero Dawkins se retira aún más, aparentemente ante la presencia de adversarios anónimos y no citados bibliográficamente:

> Mi impresión personal es que su principal valor (el meme del meme) puede radicar no tanto en ayudarnos a comprender la cultura humana como en aguzar nuestra percepción de la selección natural genética. Esta es la única razón que me hace ser lo suficientemente presuntuoso para discutir, ya que no conozco lo suficiente la literatura existente sobre la cultura humana para hacer una contribución autorizada sobre este tema (*ibidem*: 112).

Sugiero que el punto de vista del meme sobre lo que sucedió al meme del meme es completamente obvio: las mentes «humanistas» han dispuesto un conjunto de filtros especialmente agresivos contra los memes provenientes de la «sociobiología», y una vez que Dawkins fue identificado como sociobiólogo, este dato casi garantizaba el rechazo de cualquier cosa que este intruso dijera sobre la cultura, no por buenas razones sino sencillamente por un cierto tipo de rechazo inmunológico[128].

Podemos ver por qué ha sido así. La perspectiva desde el ojo del meme es un reto para uno de los axiomas centrales de las humanidades. Dawkins señala que en nuestras explicaciones tendemos a pasar por alto el hecho fundamental de que «una característica cultural distintiva puede haber evolucionado en el sentido de que simplemente es *ventajoso para ella misma*» (1976:214). Esta es una nueva forma de pensar acerca de las ideas, pero ¿es un buen método? Cuando hayamos respondido a esta pregunta, sabremos si el meme del meme es algo que estamos en condiciones o no de explotar y replicar.

La primera regla de los memes, como sucedió con los genes, es que la replicación no es necesariamente para el bien de algo; ¡surgen replicadores que son buenos para… replicarse, por cualquier razón!

Un meme que hace que sus cuerpos corran sobre desfiladeros, por ejemplo, tendrá la suerte de un gen que hace cuerpos que corran sobre desfiladeros. Tenderán a ser eliminados del *pool* de memes… Pero esto no significa que el último criterio para el éxito en la selección del meme sea la supervivencia del gen… Obviamente un meme que hace que los individuos que lo poseen se maten a sí mismos tiene una grave desventaja, pero no es necesariamente fatal… Una mente suicida puede extenderse, como cuando un dramático y bien intencionado martirio inspira a otros a morir por un apasionado amor, y éste a su vez, inspira a otros a morir, y así sucesivamente (Dawkins, 1982:110-111).

La publicidad sobre la cual habló Dawkins es crucial y tiene un paralelismo directo en la medicina darwiniana. Como Williams y Nesse [1991] han señalado, los organismos causantes de las enfermedades (parásitos, bacterias y virus) dependen, para su supervivencia a largo plazo, de saltar de huésped en huésped, y esto conlleva importantes implicaciones. Dependiendo de cómo se han extendido —a través de un estornudo o mediante contacto sexual, por ejemplo, más que por la vía de un mosquito que pica primero en una persona infectada y después en una no infectada— su futuro depende de que su huésped se mantenga en pie, y no en su lecho de muerte. Variantes más benignas serán favorecidas por la selección natural si las condiciones para la replicación de los organismos pueden ajustarse de tal manera que el no dañar a sus huéspedes sea en su propio interés. Por la misma razón, podemos

observar que, en igualdad de condiciones, memes benignos y no dañinos tenderán a proliferar y aquellos que tienden a ser fatales para las mentes que los transportan sólo pueden proliferar si tienen alguna forma de «hacerse publicidad a sí mismos» antes —o mientras— se encuentran en el mismo barco. Supongamos que James Jones (el líder de la secta The Peoples Temple que se suicidó colectivamente en la Guayana en 1978) encuentra o se imagina un argumento verdaderamente apremiante en favor del suicidio, tan apremiante que lo lleva a suicidarse. Si no deja una nota explicando por qué lo ha hecho, el meme en cuestión —al menos el que pertenece al linaje de Jones— no se extenderá.

El punto más importante que señala Dawkins es que no existe la *necesaria* conexión entre el poder replicativo del meme, su aptitud desde *su* punto de vista y su contribución a *nuestra* aptitud, sea cual sea el criterio con que la juzgamos. Esta es una observación inquietante, aunque la situación no es totalmente desesperada. Mientras algunos mentes nos manipulan definitivamente para que colaboremos en su replicación, a pesar de que los consideramos inútiles o feos, e incluso peligrosos para nuestra salud y bienestar, muchos —la mayoría, si somos afortunados— de los memes que se replican a sí mismos lo hacen no precisamente con nuestra bendición sino *a causa de* nuestra estima por ellos. Creo que existe poca controversia sobre el hecho de que los siguientes memes son, considerando todas las cosas, buenos desde nuestra perspectiva y no precisamente desde su propia perspectiva como autorreplicadores egoístas: memes muy generales como cooperación, música, escritura, calendarios, educación, conciencia ambiental, reducción de armamentos; y memes más específicos como el *Dilema del Prisionero, Las Bodas de Fígaro, Moby Dick*, las botellas reciclables y los acuerdos SALT. Otros memes son más controvertidos; observamos por qué se extienden y por qué, a fin de cuentas, los toleramos, a pesar de los problemas que nos causan: coloración de las películas clásicas, anuncios en televisión, el ideal de la corrección política. Incluso otros son perniciosos, pero extraordinariamente difíciles de erradicar: antisemitismo, secuestros de aviones, botes de spray para pintadas, virus de ordenadores[129].

Nuestra visión normal de las ideas es también una versión normativa: incorpora un canon o un ideal acerca de la idea que debemos aceptar, admirar o aprobar. Dicho brevemente, debemos aceptar lo verdadero y lo bello. De acuerdo con la visión normal, las siguientes son tautologías virtuales, verdades triviales que no merecen que se gaste la tinta escribiéndolas: «La gente creía en la idea X debido a que X era considerada verdadera», «La

gente aprobó X debido a que la gente encontró que X era bella». Estas normas no son totalmente obvias sino *constitutivas*: plantean las reglas por medio de las cuales pensamos acerca de las ideas. Necesitamos explicaciones tan sólo cuando hay desviaciones de estas normas. Nadie tiene que explicar por qué un libro intenta estar lleno de sentencias *verdaderas* ni por qué un artista puede esforzarse en hacer algo bello; «esto es justamente lo esperado». El estatus constitutivo de estas normas da fundamento al aire de paradoja en aberraciones tales como «El Museo Metropolitano de las Banalidades» o «La Enciclopedia de las Falsedades». Lo que requiere especial explicación, desde el punto de vista de la normalidad, son los casos en los que una idea no es aceptada, a pesar de su verdad o su belleza o, por el contrario, sí lo es a pesar de su fealdad o falsedad.

El punto de vista del meme da a entender que se trata de una alternativa a esta perspectiva normal. Lo que es tautológico para esta perspectiva es: «El meme X se extiende entre la gente debido a que X es un buen replicador». En la física se encuentra un bonito paralelismo. La física de Aristóteles supone que un objeto en movimiento continuo para moverse en línea recta requiere una explicación, respecto a qué fuerzas actúan sobre él continuamente. En posición dominante respecto a la gran desviación que supuso la gran perspectiva de Newton estaba la idea de que dicho movimiento rectilíneo no necesitaba explicación; sólo las desviaciones de éste, es decir, sus aceleraciones. Un paralelismo mejor puede encontrarse en la biología. Antes de que Williams y Dawkins apuntarán a la perspectiva del gen, como alternativa, los teóricos de la evolución tendían a pensar que era completamente *obvio* que las adaptaciones existiesen porque eran buenas para los organismos. Ahora conocemos mejor este problema. La perspectiva centrada en el gen es valiosa, precisamente debido a que se ocupa de casos «excepcionales», en los cuales el bien del organismo no cuenta nada, y muestra cómo la circunstancia «normal» es una regularidad derivativa y excepcional, no una verdad de pura razón, como parecía ser desde la vieja perspectiva.

Las perspectivas para la teoría de los memes llegan a ser interesantes tan sólo cuando observamos las excepciones, las circunstancias bajo las cuales hay una separación de las dos perspectivas. Para que la teoría de los memes tenga alguna garantía de ser aceptada ha de permitirnos comprender mejor las desviaciones del esquema normal. (Nótese que, en sus propios términos, si el meme del meme se replica o no con éxito es estrictamente independiente de

su virtud epistemológica; puede extenderse a pesar de su carácter pernicioso, y extinguirse a pesar de sus virtudes).

Afortunadamente para nosotros existe una correlación no aleatoria entre las dos perspectivas, justamente como sucede entre lo que es bueno para la General Motors y lo que es bueno para Norteamérica. No es un accidente que los memes que se replican tiendan a ser buenos para nosotros, no para nuestra aptitud *biológica* (el sardónico comentario de Williams sobre los feligreses es absolutamente correcto en este sentido), sino para cualquier cosa que estimemos[130]. Y nunca olvidemos el punto crucial: los hechos que estimamos, cualesquiera que sean —nuestros más altos valores—, son, en gran medida, un producto de los memes que se han extendido con mayor éxito. Podemos albergar el deseo de proclamar que somos los responsables de nuestro *summum bonum* pero esto es una insensatez mística, a menos que admitamos que lo que somos (y en consecuencia lo que podemos persuadirnos a nosotros mismos de considerar que es el *summum bonum*) es, en sí mismo, algo que hemos aprendido a ser, sobrepasando ampliamente nuestra herencia animal. La biología ha establecido algunas limitaciones sobre lo que podemos valer: a largo plazo no sobreviviríamos, a menos que tuviéramos un hábito mejor que el del azar para escoger los memes que nos ayuden, *pero todavía no hemos visto el largo plazo*. En este planeta, el experimento de la madre naturaleza con la cultura sólo se ha desarrollado durante unos miles de generaciones. Sin embargo, tenemos buenas razones para creer que nuestros sistemas inmunológicos contra los memes no están desprovistos de esperanza, incluso si no se comportan con impericia. Podemos confiar, por sentido común, en la coincidencia de dos perspectivas: de una manera general, los buenos memes —buenos según *nuestros* criterios — tenderán a ser aquellos que son también buenos replicadores.

El puerto al que todos los memes desean llegar es la mente humana, pero una mente humana es, en sí misma, un artefacto creado cuando los memes reestructuran un cerebro humano para convertirlo en el mejor hábitat posible para ellos. Las vías de entrada y de salida se modifican para adaptarlas a las condiciones locales, y son reforzadas por varios dispositivos artificiales para potenciar la fidelidad y la prolijidad de la replicación: las mentes chinas nativas difieren claramente de las mentes nativas francesas, y las mentes que saben leer y escribir difieren de las mentes iletradas. Lo que los memes aportan, a cambio, a los organismos en los cuales residen es un incalculable cúmulo de ventajas, con algunos caballos de Troya introducidos. Los cerebros humanos normales no son todos iguales; varían considerablemente en tamaño,

forma y en miles de detalles de conexión de los cuales depende su habilidad. Pero las diferencias más notables en la habilidad humana dependen de diferencias microestructurales (aún inescrutables para la neurociencia) inducidas por los diversos memes que han penetrado en los cerebros donde han fijado su residencia. Los memes potencian las oportunidades de otros memes: el meme para la educación, por ejemplo, es un meme que refuerza el mismo proceso de la implantación del meme.

Pero si es verdad que las mentes humanas son, en gran medida, la creación de memes, entonces no podemos sostener la polaridad de la visión considerada anteriormente; es decir, no puede decirse los «memes frente a nosotros», ya que la más temprana infestación de memes ha desempeñado un papel importante en determinar quiénes o qué somos. La idea de la mente «independiente» luchando por protegerse a sí misma de memes extraños y peligrosos es un mito. Existe una persistente tensión entre el imperativo biológico de nuestros genes, por un lado, y el imperativo cultural de nuestros memes por otro, pero estaríamos locos si nos alineáramos con nuestros genes; sería cometer el más egregio error de la sociobiología *pop*. Además, como ya hemos señalado, lo que nos hace especiales es que nosotros, únicos entre todas las especies, podemos elevarnos sobre los imperativos de nuestros genes, gracias a la elevación de las grúas de nuestros memes.

Entonces, ¿sobre qué fundamento podemos apoyarnos cuando luchamos para mantenernos en pie, en plena tormenta de memes? Si la replicación puede resultar incorrecta, ¿cuál ha de ser el eterno ideal respecto al cual «nosotros» juzgaremos el valor de nuestros memes? Deberíamos observar que los memes para los conceptos normativos —*deber, bondad, verdad y belleza* — se encuentran entre los ciudadanos más protegidos de nuestras mentes; tienen un papel central entre los memes que nos constituyen. Nuestra existencia como *nosotros*, como lo que somos en tanto que seres pensantes —no lo que somos como organismos— no es independiente de estos memes.

Dawkins termina su libro *El gen egoísta* (1976) con un pasaje que probablemente muchos de sus críticos no han leído o no han comprendido:

> Tenemos el poder de desafiar a los genes egoístas de nuestro nacimiento y, si es necesario, a los genes egoístas de nuestro adoctrinamiento… Hemos sido construidos como máquinas de genes y cultivados como máquinas de memes, pero tenemos el poder de revolvernos contra nuestros creadores. Nosotros somos los únicos en la Tierra que podemos rebelarnos contra la tiranía de los replicadores egoístas.

Al distanciarse, aunque forzadamente, de las hipersimplificaciones de la sociobiología *pop*, de algún modo Dawkins exagera esta cuestión. Ese «nosotros» que trasciende no solamente sus creadores genéticos sino también

sus creadores meméticos es, como hemos visto, un mito. El mismo Dawkins reconocía esto en su último trabajo. En su libro *The Extended Phenotype* (1982) argumenta desde la perspectiva biológica que reconoce la presa del castor, la red de la araña y el nido del pájaro no como meros productos del fenotipo —el organismo individual considerado como un todo funcional— sino como partes del fenotipo, al mismo nivel de los dientes del castor, las patas de la araña y las alas del pájaro. Desde esta perspectiva, la enorme red protectora de mentes que hilamos es parte tan integral de nuestro fenotipo —para explicar nuestras competencias, nuestras oportunidades, nuestras vicisitudes— como cualquier otra dentro de nuestra más estrecha dotación biológica (esta tesis se desarrolla con mayor detalle en Dennett 1991a). No hay una discontinuidad radical; se puede ser un mamífero, un padre, un ciudadano, un erudito, un demócrata y un profesor numerario. Del mismo modo que un granero hecho por el hombre es parte integral de la ecología de la golondrina de los graneros, así las catedrales y las universidades —y las fábricas y las prisiones— son parte integral de nuestra ecología, igual que los memes sin los cuales no podríamos vivir en estos ambientes.

Pero si no soy nada por encima y por debajo de algún complejo sistema de interacciones entre mi cuerpo y los memes que lo infectan, ¿qué pasa con la responsabilidad personal? ¿Cómo puedo ser declarado responsable de mis delitos o bien honrado por mis triunfos, si no soy el capitán de mi barco? ¿Dónde está la autonomía que necesito para actuar con libre albedrío?

«Autonomía» es precisamente un término de moda en lugar de «autocontrol». Cuando la nave espacial *Viking* estaba demasiado lejos de la Tierra para ser controlada por los ingenieros de Houston, éstos le enviaron un nuevo programa, el cual eliminó el control remoto y la puso bajo autocontrol local (Dennett 1984:55). Esta maniobra la convirtió en autónoma y aunque los objetivos seguían siendo los que Houston había instalado desde su nacimiento, la nave ahora era la única responsable de tomar las decisiones en la persecución de estos objetivos. Imaginemos ahora que la nave toma tierra en algún lejano planeta habitado por hombrecitos verdes que la invaden enseguida, que manipulan su *software* y lo modifican para sus propios propósitos: convertirlo en un vehículo de recreo, por ejemplo, o bien en una guardería para sus hijos. La autonomía de la *Viking* se habría perdido cuando cayeron bajo el control de estos controladores alienígenas. Traspasar las responsabilidades de mis genes a mis memes puede parecer un paso similar, no comprometido en el camino hacia el libre albedrío. ¿Hemos roto la tiranía

de los genes egoístas solamente para caer bajo la tiranía de los egoístas memes?

Veamos lo que sucede con los simbiontes. Los parásitos son, por definición, aquellos simbiontes que resultan deletéreos para la aptitud del huésped. Consideremos el ejemplo más obvio de un meme: el meme del celibato (castidad, puedo añadir, para cerrar una notoria excusa). Este complejo meme habita el cerebro de muchos sacerdotes y monjas. Desde el punto de vista de la biología evolucionista, este complejo es deletéreo, por definición, para la aptitud: todo aquello que garantice virtualmente que la línea germinal del huésped será un callejón sin salida, sin nada que lo prolongue, rebaja la aptitud. Pero ¿cuál es la razón?, podría interpelarnos un sacerdote. «¡Yo no deseo tener hijos!». Exactamente. Pero le podemos decir: su cuerpo sí. El sacerdote se ha distanciado, de algún modo, de su propio cuerpo, lo que le ocasiona a veces problemas de autocontrol. ¿Qué hace este *propio yo* o *ego* frente a los divergentes objetivos con los que ha sido constituido? No conocemos la historia detallada de la infestación. Los jesuitas dicen con una frase famosa: «Dadme los primeros cinco años de la vida de un niño y yo puedo poseer el resto», así que seguramente es en una etapa muy precoz de la vida de un sacerdote cuando este específico meme asalta la fortaleza. O bien puede haber sido más tarde o acaso muy gradualmente. Pero cuando quiera y como quiera que esto sucediera, ha sido incorporado por el sacerdote —al menos por lo pronto— dentro de su identidad.

Yo *no* estoy diciendo que la causa de que el cuerpo del sacerdote esté «predestinado» a no tener descendencia sea una cosa mala o «no natural». Eso sería alinearme con nuestros genes egoístas, que es precisamente lo que no deseo. Estoy diciendo que éste es precisamente el más extremo y, en consecuencia, clarísimo ejemplo del proceso que nos ha hecho a todos nosotros: nuestros *propios yos* han sido creados a partir de la interacción de memes explotando y redirigiendo la maquinaria que la madre naturaleza nos ha dado. Mi cerebro está habitado por los memes para el celibato y la castidad (yo no podría escribir acerca de esta cuestión si no fuera así), pero nunca han tratado de ocupar en mí el asiento del conductor. Yo no me *identifico* con ellos. Mi cerebro también está habitado por los memes del ayuno o de la dieta, y desearía que estuvieran más a menudo en el asiento del conductor (así podría seguir una dieta sinceramente) pero, por una u otra razón, las condiciones de los memes que incorporarían el meme de la dieta que yo seguiría sinceramente, raras veces consiguen formar un gobierno estable. Ningún meme controla solitariamente a nadie; lo que hace que una persona

sea lo que él o ella es, son coaliciones de memes que gobiernan, que a largo plazo desempeñan los papeles determinando qué decisiones han de tomarse a lo largo del camino. (Revisaremos más detalladamente esta idea en los capítulos 16 y 17).

Sea o no posible que la perspectiva de los memes se convierta en ciencia, con su apariencia filosófica se ha hecho ya bastante más bien que mal, contrariamente a lo que Gould sostenía, aunque, como hemos visto, puede haber otras aplicaciones del pensamiento darwiniano en las ciencias sociales que merezcan verdaderamente la condena de Gould. ¿Cuál es, de hecho, la alternativa a esta visión darwiniana de la mente? Una última esperanza, para los que temen a Darwin, es simplemente negar que lo que sucede a los memes cuando entran en la mente pueda ser nunca explicado en términos «reduccionistas», mecánicos. Una vía sería abogar completamente por el dualismo cartesiano: la mente no puede ser precisamente el cerebro, sino, más bien, algún *otro* lugar en el cual ocurren procesos alquímicos grandes y misteriosos, transformando las materias primas que los alimentan —los objetos culturales que estamos llamando memes— en nuevos objetos que trascienden sus fuentes en caminos que están simplemente más allá del alcance de la comprensión de la ciencia[131].

Un modo ligeramente menos radical de apoyar el mismo punto de vista defensivo será conceder que la mente es, después de todo, justamente el cerebro, el cual es una entidad física ligada con la totalidad de las leyes de la física y de la química, aunque insistiendo en que, no obstante, hace su trabajo de un modo que desafía los análisis científicos. Esta vía ha sido a menudo sugerida por el lingüista Noam Chomsky y defendida con entusiasmo por su anterior colega, el filósofo y psicólogo Jerry Fodor [1983], y más recientemente por otro filósofo, Colin McGinn [1991]. Podemos ver que este punto de vista de la mente implica la tesis de los saltos evolutivos, ya que postula grandes saltos hacia delante en el espacio de diseño que quedan «explicados» como actos de traslúcidos genios o de una creatividad intrínseca o de otra cosa que desafía a la ciencia. Se insiste en que, de algún modo, el cerebro es un gancho celeste y se rehúsa aceptar simplemente lo que el taimado darwinismo ofrece: el cerebro, gracias a todas las grúas que lo han formado en primer lugar y a todas las grúas que han entrado en él en una segunda fase es, en sí mismo, un prodigioso, pero no misterioso, elevador en el espacio de diseño.

Nos costará todavía más trabajo convertir, en el capítulo 13, esta confrontación extraordinariamente metafórica en una confrontación más

literal, y resolverla. Afortunadamente para mí, gran parte de este trabajo lo he hecho previamente, así que puedo de nuevo evitar reinventar la rueda simplemente reutilizando una rueda que ya he hecho antes. Mi próxima readaptación procede de mi Darwin Lecture pronunciada en 1991, en el Darwin College, Cambridge (Dennett 1994b).

CAPÍTULO 13

Perdiendo nuestras mentes en beneficio de Darwin

I. *El papel del lenguaje en la inteligencia*

> Cuando las ideas fallan, las palabras resultan muy útiles.
>
> Anónimo[132]

Nosotros no somos como los otros animales; nuestras mentes nos han ido separando de ellos. Esta es la afirmación que inspira tan apasionada defensa. Es curioso que la gente que tanto desea defender esta diferencia tenga que ser tan reacia a examinar la evidencia a su favor que procede de la biología evolucionista, la etología, la primatología y la ciencia cognitiva. Temen, presumiblemente, que puedan aprender que, aunque somos diferentes, no lo somos lo *bastante* como para alcanzar esa diferencia incardinada en la propia vida, que tanto estiman. Para Descartes, después de todo, la diferencia era absoluta y metafísica: los animales eran autómatas insensatos; *nosotros* tenemos almas. Descartes y sus seguidores han soportado durante siglos ataques calumniosos de los amantes de los animales que deploraban su afirmación de que éstos no tenían almas. Los críticos con mentalidad más teórica, desde el polo opuesto, han deplorado la pusilanimidad de Descartes: ¿cómo podía un mecanicista tan ingenioso y con tan buenos fundamentos acobardarse hasta el punto de hacer una excepción con la humanidad? Nuestras mentes son *naturalmente* nuestros cerebros y, en consecuencia, son, en último término, «máquinas» estupendas y complejas; la diferencia entre nosotros y el resto de los animales no es de tipo metafísico, sino de mayor entidad. Como he demostrado, no es una coincidencia que aquellos que deploran la inteligencia artificial sean los mismos que deploran la versión evolucionista de la mentalidad humana: si las mentes humanas son productos no milagrosos de la evolución, entonces estas mentes son, necesariamente, artefactos y todos sus poderes deben tener, finalmente, una explicación

«mecánica». Nosotros somos descendientes de macros y hechos de macros, y nada de lo que podamos hacer está más allá del poder dimanado de enormes asambleas de macros (asambleas en el espacio y en el tiempo).

Además, existe una enorme diferencia entre nuestras mentes y las de otras especies, un enorme golfo lo suficientemente amplio como para llegar a constituirse en una diferencia moral. Esta diferencia se debe a dos factores entremezclados, cada uno de los cuales requiere una explicación darwiniana: 1) el cerebro con el que hemos nacido tiene características que están ausentes en otros cerebros, características que han evolucionado bajo la presión de la selección durante los últimos seis millones de años, aproximadamente, y 2) estas características hacen posible una enorme elaboración de poderes que se acrecientan al compartir, a través de la transmisión cultural, la riqueza del diseño. El lenguaje es el fenómeno clave, verdadero pivote que une estos dos factores. Nosotros, los seres humanos, podemos no ser la más admirable de las especies que viven en este planeta, o la que es más probable que sobreviva otro milenio, pero somos, sin duda, la más inteligente de todas. Somos también la única especie con lenguaje.

¿Es esto verdad? ¿No tienen las ballenas y los delfines, los monos «vervet» y las abejas (la lista puede ampliarse) *algo parecido* a un lenguaje? ¿No se les ha enseñado a los chimpancés en los laboratorios *algo parecido* a un lenguaje rudimentario? Sí, y el lenguaje corporal es algo parecido a un lenguaje, y la música es un tipo de lenguaje internacional y la política es una forma de lenguaje, y el complejo mundo de los olores y del olfato es otro lenguaje, lenguaje éste lleno de carga emocional, etc. A veces parece que la mayor alabanza que podemos aplicar a un fenómeno en estudio es la afirmación de que sus complejidades justifican que sea considerado, de algún modo, como un lenguaje. Esta admiración por el lenguaje —por el lenguaje real, el tipo de lenguaje que sólo usa el ser humano— está bien fundamentada. Las propiedades expresivas del lenguaje real, codificadoras de información, prácticamente no tienen límites (al menos en alguna dimensión) y los poderes que otras especies adquieren en virtud de usar protolenguajes, hemi-semi-demi-lenguajes, son naturalmente similares a los poderes que adquirimos gracias a nuestro uso del lenguaje real. Estas otras especies suben unos cuantos escalones de la montaña en cuya cima, gracias al lenguaje, residimos nosotros los humanos. Observar las enormes diferencias entre sus ganancias y las nuestras es una vía de aproximación a la pregunta cuya respuesta vamos a intentar: ¿cómo contribuye el lenguaje a la inteligencia?

¿Qué variedades de pensamiento requiere el lenguaje? ¿Qué variedades de pensamiento (si hay alguna) son posibles sin el lenguaje? Observemos a una chimpancé con su cara sentimental, su mirada inquisitiva y sus hábiles dedos, y tendremos la impresión, casi definitiva, de que allí dentro existe una mente, pero mientras más la observamos, más se difumina, ante nuestros ojos, la idea de esa mente. En cierto modo, la chimpancé parece muy humana, llena de perspicacia; pero muy pronto nos damos cuenta (para nuestra decepción o alivio, según nuestras esperanzas) de que, desde otros puntos de vista, la chimpancé es muy estúpida, muy falta de comprensión y se encuentra muy alejada, a una distancia inalcanzable, del mundo de los humanos. ¿Cómo puede un chimpancé que tan obviamente comprende A, fallar en la comprensión de B) Consideremos varias simples cuestiones acerca de los chimpancés. ¿Podrían aprender a dominar un fuego, es decir, a reunir la leña, mantenerla seca, guardar el carbón, trocear la madera, mantener el tamaño de la hoguera dentro de sus límites apropiados? Y si no pudieran inventar por sí mismos estas nuevas actividades, ¿podrían ser entrenados por seres humanos para hacer estas cosas? Aquí se nos plantea otra cuestión. Supongamos que imaginamos algo nuevo; por ejemplo, les invito a imaginarse a un hombre que asciende por una cuerda con un cubo de plástico para la basura colocado sobre su cabeza. Una tarea mental que es fácil para nosotros. ¿Podría un chimpancé hacer lo mismo en su mente? Me extrañaría. Yo escojo los elementos —hombre, cuerda, ascenso, cubo, cabeza— como objetos familiares en el mundo de la percepción y de la conducta de un chimpancé que vive en un laboratorio, pero me pregunto si un chimpancé podría reunir todos estos elementos en esta nueva forma, aunque fuera por accidente. Vosotros —lectores— habéis sido incitados a realizar este acto mental por mi sugerencia verbal, y probablemente a menudo realizáis actos mentales similares como respuesta propia a sugerencias verbales que os dais a vosotros mismos, no ya en voz baja, sino incluso claramente con palabras. ¿Podría ser de otra manera? ¿Podría un chimpancé realizar dicho acto mental sin la ayuda de una sugerencia verbal?

Estas son preguntas bastante simples acerca de los chimpancés, aunque nadie conoce todavía las respuestas. Las respuestas no son imposibles de obtener, aunque no es fácil; experimentos controlados pueden generar las respuestas, las cuales podrían arrojar luz sobre el papel del lenguaje en la conversión de cerebros en mentes, como ha sucedido en nosotros los humanos. Planteo aquí la cuestión de si los chimpancés pueden aprender a manejar un fuego debido a que, en algún momento de la prehistoria, nuestros

ancestros dominaron el fuego. ¿Fue necesario el lenguaje para este gran avance en la civilización? Algunas evidencias sugieren que esto sucedió hace cientos de miles de años —o incluso hace un millón de años (Donald 1991:114)— *antes* de la introducción del lenguaje, aunque, naturalmente, después de que nuestra línea de homínidos se separara de los antecesores de los simios modernos, como los chimpancés. Las opiniones difieren profundamente. Muchos investigadores están convencidos de que el lenguaje comenzó mucho antes, con suficiente tiempo para garantizar el dominio del fuego (Pinker 1994). Se puede incluso argumentar que el dominio del fuego es, en sí mismo, incontrovertible evidencia de la existencia de un lenguaje precoz, si aceptamos que esa proeza mental requirió un lenguaje rudimentario. ¿O es que el manejo del fuego no es un asunto tan importante? Quizá la única razón por la que no encontramos a los chimpancés sentados en la selva alrededor del fuego de una hoguera es que en sus lluviosos hábitats no hay nunca bastante leña seca a su disposición para que un fuego pueda ser dominado. (A los chimpancés pigmeos de Sue Savage-Rumbaugh, en Atlanta, les gusta ir de *picnic* a los bosques y disfrutan mirando fijamente las llamas de las fogatas, como hacemos nosotros, pero Sue me ha dicho que duda que se les pueda confiar el cuidado de un fuego, aunque hayan sido entrenados).

Si las termitas pueden crear ciudades de barro muy elaboradas y bien ventiladas y el pájaro tejedor puede tejer nidos colgantes de una audaz ingeniería, y los castores construir presas que tardan meses en terminar, ¿no podrían los chimpancés cuidar un fuego? Esta cuestión retórica se plantea subiendo por una equivocada escalera de habilidades. Ignora la posibilidad, bien demostrada, y de modo independiente, de que hay dos métodos muy diferentes para construir presas: el método que utilizan los castores y el método que utilizamos los humanos. Las diferencias no se encuentran necesariamente en los productos sino en las estructuras de control dentro de los cerebros que las han creado. Un niño puede estudiar cómo el pájaro tejedor construye su nido y luego imitarlo, encontrando los apropiados fragmentos de hierbas y tejiéndolos en el orden adecuado, hasta llegar a construir, siguiendo una idéntica serie de pasos, un nido idéntico. Una película de los dos procesos de construcción que se desarrollaran paralelamente podría abrumarnos con la idea de que estamos viendo dos veces el mismo fenómeno, pero sería un gran error imputar al pájaro el tipo de procesos de pensamiento que conocemos o imaginamos que se producen en el niño. Puede haber muy poco en común entre los procesos que tienen lugar en el cerebro del niño y en el cerebro del pájaro. El pájaro está (aparentemente)

dotado con una colección de subrutinas minimalistas integradas con un propósito específico, bien diseñadas por la evolución de acuerdo con el bien conocido principio del espionaje que dice que cada agente debe *saber sólo lo necesario*: dar a cada agente tan sólo la mínima información necesaria para cumplir su participación en la misión encomendada.

Los sistemas de control diseñados bajo este principio pueden ser sorprendentemente exitosos —prueba de ello son los nidos de los pájaros— dondequiera que el medio ambiente posea suficiente simplicidad y regularidad y, en consecuencia, sea lo suficientemente predecible como para favorecer el prediseño de la totalidad del sistema. De hecho, el propio diseño del sistema establece la predicción —en realidad, una apuesta— de que el medio ambiente será el adecuado para que el sistema funcione. Sin embargo, cuando se incrementa la complejidad del medio ambiente encontrado y la incapacidad de predicción llega a ser un problema más grave, se cuenta con un principio de diseño diferente: el *principio del equipo o comando*, ilustrado en películas como *Los cañones de Navarone*: proporcionar a cada agente todos los conocimientos que sea posible sobre el proyecto total, de modo que el equipo pueda actuar sin limitaciones cuando surjan obstáculos imprevistos.

De este modo, se establece una especie de divisoria de aguas en el terreno del espacio de diseño evolutivo; cuando un problema de control se cruza en esta divisoria, puede ser un problema de suerte en qué dirección empujará la evolución a los descendientes exitosos. En este sentido, quizás haya, aproximadamente, dos vías para dominar las hogueras: la vía seguida por el castor constructor de presas y la nuestra. Si es así, para nosotros fue una buena cosa que nuestros ancestros no escogieran la vía del castor, porque si lo hubieran hecho los bosques hoy podrían estar llenos de simios sentados alrededor de fogatas, pero nosotros no estaríamos aquí para maravillarlos.

Propongo una estructura en la cual podamos situar las diversas opciones de diseño para los cerebros, con el fin de averiguar de dónde viene su poder. Es una estructura excesivamente hipersimplificada, pero la idealización es el precio que se debe pagar a menudo por las visiones sinópticas. Yo la llamo la *torre de generación y prueba*; cada vez que se construye una nueva planta de la torre, eleva al organismo a este nivel, con el fin de encontrar mejores movimientos, y encontrarlos con mayor eficiencia**[133]**.

En el principio —una vez que la bomba de Darwin había sido cebada— estaba la evolución darwiniana de las especies por selección natural. Variados organismos candidatos se generaron ciegamente a través de procesos más o menos arbitrarios de recombinación y mutación de genes. Estos organismos

se sometieron a prueba en el terreno y sólo los mejores diseños sobrevivieron. Este es el primer piso de la torre. Permítaseme llamar a los habitantes de este nivel *criaturas darwinianas*.

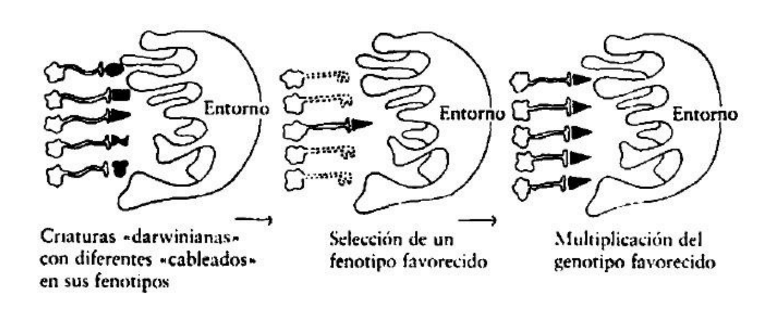


Figura 13.1

Este proceso se desarrolló a lo largo de muchos millones de ciclos, produciendo muchos y admirables diseños, tanto de plantas como de animales, y, eventualmente, entre sus nuevas creaciones había algunos diseños con la propiedad denominada plasticidad fenotípica. Los organismos individuales, candidatos a ser seleccionados, no estaban totalmente diseñados en el momento de su nacimiento o, en otras palabras, había elementos en sus diseños que podían ajustarse, según los acontecimientos que ocurrieran al ser probados en el terreno. (Esto es lo que hace posible el efecto Baldwin, tal como vimos en el capítulo 3, aunque ahora vamos a dirigir nuestra atención al diseño del interior de los organismos que pone en pie esta grúa). Podemos suponer que algunos de estos organismos candidatos no estaban en mejores condiciones que sus primos dotados de un cableado para circuitos incorporados, dado que no disponían de un modo de favorecer (seleccionándolas para ser repetidas) las opciones de conducta de las que estaban equipados «para someterlas a prueba», pero otros, también podemos suponerlo, fueron bastante afortunados por tener «reforzadores» en circuitos incorporados, los cuales favorecerían los movimientos inteligentes, es decir, aquellas acciones que eran mejores para sus agentes. Estos individuos se enfrentarían con el medio ambiente generando una variedad de acciones, las cuales fueron probando, una por una, hasta encontrar la que funcionara. A este subgrupo de criaturas darwinianas, unas criaturas con plasticidad condicionable, podemos denominarlo el de las *criaturas skinnerianas*, dado que, como B. F. Skinner se enorgullecía de haber demostrado, los condicionamientos del operante no eran precisamente análogos a los de la selección natural darwiniana, aunque exista una continuidad entre ambas.

«Allí donde se interrumpen conductas heredadas, toma el control la modificabilidad heredada del proceso de condicionamiento» (Skinner 1953:83).

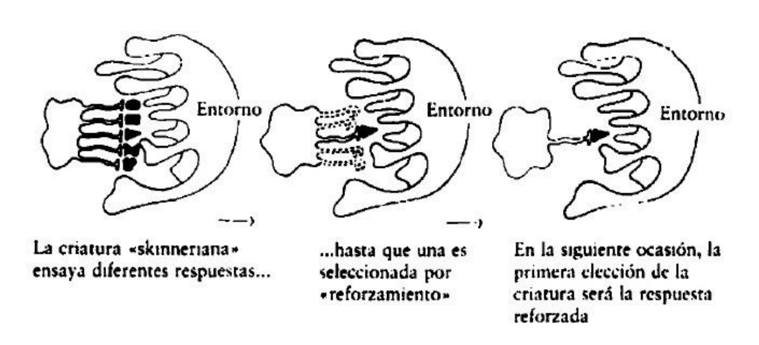


Figura 13.2

Poseer la capacidad del condicionamiento skinneriano es excelente, en tanto en cuanto uno no sea antes eliminado a causa de alguno de sus tempranos errores. Un sistema mejor que el condicionamiento skinneriano exige una *selección previa* llevada a cabo entre todas las posibles conductas o acciones, para eliminar las opciones verdaderamente estúpidas antes de arriesgarlas en el áspero mundo. Nosotros, los seres humanos, somos criaturas capaces de un tercer refinamiento, aunque no en solitario. A los beneficiarios de este tercer piso de nuestra torre los podemos calificar de *criaturas popperiana*, dado que, como en cierta ocasión escribió sir Karl Popper con elegancia, esta potenciación del diseño «permite que nuestras hipótesis mueran, ocupando nuestro lugar». Al contrario de las que son simplemente criaturas skinnerianas, la mayoría de las cuales sobreviven sólo gracias a que hicieron los primeros movimientos afortunados, las criaturas popperianas sobreviven gracias a que son lo bastante inteligentes para hacer los primeros movimientos con algo más que suerte. Naturalmente, estas criaturas son afortunadas por ser inteligentes, lo que es mejor que ser solamente afortunadas.

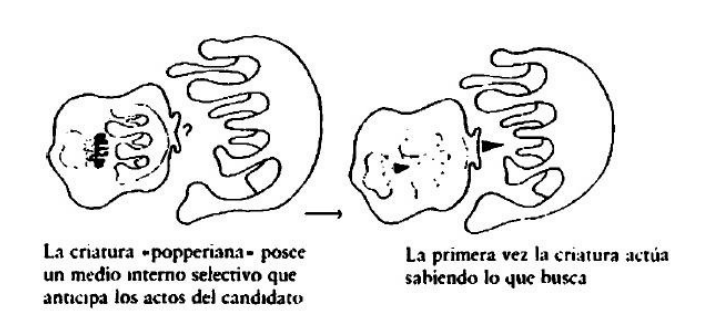


Figura 13.3

Pero ¿cómo hacen esta preselección los agentes popperíanos? ¿De dónde procede el *feedback*? Debe proceder de una especie de *medio interno* —una cosa u otra interna— estructurada de tal manera que favorece, más a menudo, las acciones subrogadas que las propias acciones que el mundo real también protege, si realmente se realizaran. En resumen, el medio interno, sea lo que sea, debe contener grandes cantidades de *información* acerca del medio externo y sus regularidades. Ninguna otra cosa —excepto la magia— puede aportar una preselección que valga la pena. No obstante, tenemos que andar con mucho cuidado por este terreno y no creer que este medio interno es una simple réplica del mundo exterior, que reproduce sus contingencias físicas. (¡Si fuera así, en ese milagroso mundo de juguete, la pequeña hornilla caliente ubicada en el interior de nuestra cabeza estaría realmente lo bastante caliente como para quemarnos el dedo que colocásemos sobre ella!). La información sobre el mundo ha de estar allí, pero también estructurada de tal modo que exista una explicación no milagrosa de cómo llegó hasta allí, cómo se mantiene y cómo consigue realmente los efectos preselectivos que son su *raison d'être*.

¿Qué animales son criaturas popperianas y cuáles simplemente criaturas skinnerianas? Las palomas eran los animales de experimentación favoritos de B. F. Skinner y él y sus seguidores desarrollaron la tecnología del condicionamiento del operante (la paloma en este caso) hasta un nivel muy elaborado, con lo que conseguían que las palomas exhibieran conductas aprendidas que eran extrañas y sofisticadas. Es notorio que Skinner y sus seguidores nunca consiguieran demostrar que las palomas *no* eran criaturas popperianas, e investigaron en numerosas y diferentes especies, desde pulpos hasta peces y mamíferos, con la firme sospecha de que si hay algunas criaturas puramente skinnerianas, capaces únicamente de aprender mediante el ciego método de la prueba y el error éstas deben encontrarse entre los

simples invertebrados. La babosa de mar *Aplysia* ha desplazado, más o menos, a la paloma del foco de atención de aquellos que estudian los mecanismos del simple condicionamiento. (Los investigadores, sin dudas y sin controversias, ordenan las especies según lo inteligentes que éstas sean. Esto implica no un respaldo miópico a la gran cadena del ser, ni tampoco un respaldo a las hipótesis indefendibles acerca del ascenso por la escalera del progreso. Todo está función de medidas objetivas de la competencia cognitiva. El pulpo, por ejemplo, es asombrosamente inteligente, un hecho que no nos sorprendería si no hubiera métodos para medir la inteligencia que sean independientes del chauvinismo filogeqético).

En consecuencia, no nos diferenciamos de las otras especies por ser criaturas popperianas. Ni mucho menos; mamíferos, pájaros, reptiles y peces, todos exhiben la capacidad de utilizar información de su medio ambiente para seleccionar sus opciones de conducta antes de golpear vigorosamente. Alcanzamos en este momento el piso de la torre sobre el cual deseo construir. Una vez que hemos llegado a las criaturas popperianas, criaturas cuyos cerebros poseen el potencial de ser conformados en su medio interno de modo que cuenten con habilidades preselectivas, ¿qué sucede inmediatamente después? ¿Cómo se incorpora la nueva información acerca del medio externo en estos cerebros? Es allí donde las decisiones de diseño más tempranas vuelven a estar presentes —como limitaciones— ante el diseñador. En especial, las elecciones que la evolución ha hecho ya entre el principio de saber sólo lo necesario y el principio del equipo o comando, plantean ahora mayores limitaciones a las opciones que tratan de mejorar el diseño. Si un tipo de diseño del cerebro de una especie concreta ha descendido evolutivamente según el «principio» de saber sólo lo necesario con respecto a algún problema de control, tan sólo modificaciones menores (podría decirse que ajustes de sintonía) pueden hacerse con facilidad en las estructuras existentes, de modo que la única esperanza de poder llevar a cabo una revisión mayor del medio interno, para enfrentarse con nuevos problemas, nuevas características del medio externo, es *sumergir* el viejo circuito incorporado bajo una nueva capa de control preventivo (un tema desarrollado por ejemplo en el trabajo del investigador de inteligencia artificial Rodney Brooks 1991). Son estos niveles de control más elevados los que tienen el potencial de producir enormes incrementos de versatilidad. Y es en estos niveles en particular donde debemos contemplar el papel del lenguaje (cuando entre finalmente en escena) convirtiendo *nuestros* cerebros en virtuosos preseleccionadores.

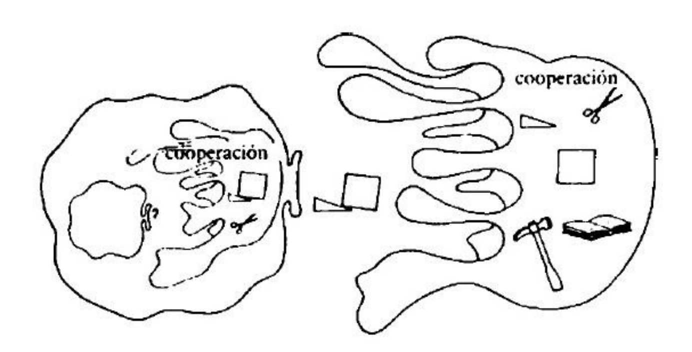
Participamos con nuestra cuota de conducta más bien rutinaria e insensata, pero nuestros actos importantes son, a menudo, dirigidos sobre el mundo con increíble habilidad, hasta componer proyectos exquisitamente diseñados bajo la influencia de vastas bibliotecas de información acerca del mundo. Las acciones instintivas que compartimos con otras especies muestran los beneficios derivados de las inquietantes exploraciones de nuestros ancestros. Las acciones imitativas que compartimos con algunos animales más elevados pueden mostrar los beneficios de la información reunida no precisamente por nuestros ancestros, sino también por nuestros grupos sociales a lo largo de las generaciones, información transmitida no genéticamente sino por una «tradición» de imitación. Pero nuestros actos más deliberadamente planeados muestran los beneficios de la información reunida y transmitida por los seres de nuestra misma especie en toda cultura, incluyendo, además, objetos de información que ningún individuo singular ha incorporado o entendido en ningún sentido. Y, aunque parte de esta información puede ser de antigua adquisición, la mayor parte es nueva. ¡Cuando se comparan las escalas temporales de la evolución genética y cultural, es útil tener presente que hoy nosotros —cada uno de nosotros— puede fácilmente comprender muchas ideas que resultaban sencillamente inconcebibles para los genios de la generación de nuestros abuelos!

Los sucesores de las meras criaturas popperianas son aquellos cuyos medios internos están informados por las porciones *diseñadas* del medio externo. Podemos denominar a este sub-sub-subgrupo de criaturas darwinianas como criaturas *gregorianas*, ya que el psicólogo británico Richard Gregory es, a mi entender, el teórico preeminente en lo tocante al papel de la información (o más exactamente, lo que Gregory llama la inteligencia potencial) en la creación de movimientos inteligentes (o lo que Gregory llama la inteligencia cinética). Gregory observa que un par de tijeras, como un artefacto bien diseñado, no es solo un resultado de la inteligencia sino algo que dota de inteligencia (inteligencia potencial externa) en un sentido muy directo e intuitivo: cuando damos a alguien un par de tijeras, incrementamos su potencial para alcanzar, con mayor rapidez y seguridad, los movimientos inteligentes (Gregory 1981:31 ISS).

Los antropólogos hace tiempo que han reconocido que la introducción del uso de las herramientas se asoció a un mayor incremento de la inteligencia. Los chimpancés en la selva pescan termitas con una especie de cañas de pescar muy rudimentarias. Este hecho tiene un significado adicional cuando aprendemos que no todos los chimpancés tienen esta habilidad o maña; en

algunas «culturas» de chimpancés las termitas son un alimento disponible, pero no ha sido explotado. Este hecho nos recuerda que el uso de herramientas es un signo de inteligencia en un doble sentido; usarlas no sólo *requiere* inteligencia para reconocer y mantener una herramienta (abandonar una y fabricar otra) sino que el uso de la herramienta *confiere* inteligencia a aquellos que son bastante afortunados para que les sea dada la herramienta. Mientras mejor diseñada está la herramienta (mientras más información se encuentra embebida en su fabricación) más inteligencia potencial confiere su uso. Y entre las herramientas preeminentes, nos recuerda Gregory, están las que llama las «herramientas de la mente», las palabras.

Las palabras y otras herramientas de la mente proporcionan a la criatura gregoriana un medio interno que le permite construir generadores de movimientos y aún más sutiles probadores de movimientos. Las criaturas skinnerianas se preguntan a sí mismas: «¿Cuál será próximo movimiento que debo ejecutar?», y no disponen de una clave para responder hasta que han recibido algunos duros golpes. Las criaturas popperianas pueden lograr un gran avance al preguntarse a sí mismas: «¿Qué he de pensar acerca del próximo movimiento?», antes de interrogarse a sí mismas: «¿Cuál será el próximo movimiento que debo hacer?». Las criaturas gregorianas dan un gran paso hacia delante aprendiendo a pensar mejor sobre lo que deben pensar a propósito del siguiente movimiento, y así en adelante, construyendo una torre de sucesivas reflexiones internas sin límite fijado o discernible.



La criatura gregoriana importa del entorno (cultural) herramientas mentales que mejoran tanto los «generadores» como los «probadores»

Figura 13.4

¿Qué le sucede a un cerebro humano o de homínido cuando llega a contener palabras? Concretamente, ¿cuál es la forma de este ambiente cuando las palabras entran en él por primera vez? Definitivamente, *no* es un campo en el que nunca se haya jugado, ni tampoco una *tabula rasa*. Nuestras recién

halladas palabras deben fijarse ellas mismas en las colinas y en los valles de un paisaje de considerable complejidad. Gracias a presiones evolutivas más tempranas, nuestros espacios de calidad innata son específicos de especies, narcisistas e incluso idiosincrásicos de individuo a individuo. Algunos investigadores están actualmente explorando partes de este terreno. El psicólogo Frank Keil [1992] y sus colegas de la Cornell University tienen evidencias de que ciertos conceptos muy abstractos, por ejemplo, los conceptos de *estar vivo* o de *propiedad*, tienen un punto de partida impuesto genéticamente en la caja de herramientas mentales del niño; cuando palabras específicas como *propiedad, dar y tomar, guardar y esconder* y otras similares entran en el cerebro de un niño, encuentran ya habitáculos parcialmente construidos para ellas. Ray Jackendoíf [1993] y otros lingüistas han identificado estructuras fundamentales de representación espacial, excelentemente diseñadas para potenciar el control de la *locomoción* y de la *colocación* de las cosas móviles, que sustentan nuestras intuiciones acerca de conceptos tales como *al lado, encima, detrás*, y similares. Nicholas Humphrey [1976, 1983, 1986] ha argumentado que debe existir una predisposición genética para adoptar la posición intencional, y Alan Leslie (1992) y otros han desarrollado evidencias a favor de esta tesis, bajo la forma de lo que llaman la «teoría del módulo mental», diseñado para generar creencias de segundo orden (creencias acerca de las creencias y otros estados mentales acerca de otros). Algunos niños autistas parecen estar bien descritos como sufriendo de una alteración de este módulo y para corregirla estos niños pueden hacer, ocasionalmente, interesantes ajustes compensatorios. (Para una revisión, consúltese a Baron-Cohen 1995). De este modo las palabras (y en consecuencia los memes) que fijan su residencia en el cerebro, así como muchas novedades más tempranas del diseño que hemos considerado, potencian y conforman estructuras preexistentes, más que generar arquitecturas enteramente nuevas (véase Sperbe 1985 para una revisión darwiniana de esta readaptación de las funciones genéticamente proporcionadas por funciones transmitidas culturalmente). Aunque estas funciones nuevamente rediseñadas no están hechas a partir de un tejido intacto, crean una nueva y explosiva capacidad para seguir adelante.

Un modelo interno permite a un sistema mirar hacia delante, a las futuras consecuencias de las acciones actuales, sin comprometerse a sí mismo, realmente, a llevar a cabo aquellas acciones. En concreto, el sistema con modelo interno puede evitar actos que serían irreparables y que conducirían a un futuro desastre (¡cayendo por un despeñadero!). No de manera tan dramática, pero igualmente importante, este modelo hace posible que el agente realice movimientos para preparar un escenario transitorio que sirva de base a otros movimientos que sean obviamente ventajosos. La verdadera esencia de una ventaja

competitiva, sea en el ajedrez o en la economía, es el descubrimiento y la ejecución de los movimientos que preparan el escenario (Holland 1992:25).

Esta es, pues, la grúa para culminar todas las grúas: un explorador que tiene en cuenta la previsión, que puede ver más allá de la inmediata vecindad de las opciones. Pero ¿hasta qué punto puede ser bueno el «escenario» sin la intervención del lenguaje para ayudar a controlar la manipulación del modelo? ¿Hasta qué punto han de ser complejos y a largo plazo los planes para el futuro? Esta es la relevancia de mi pregunta acerca de las capacidades del chimpancé para visualizar una nueva escena. Darwin estaba convencido de que el lenguaje era el prerrequisito para los «largos trenes de pensamiento» y esta afirmación ha sido diferentemente apoyada por varios teóricos recientes, especialmente Julian Jaynes [1976] y Howard Margolis [1987]. Largos trenes del pensamiento han de ser controlados o, de lo contrario, marcharían sin rumbo, siguiendo deliciosas pero fútiles ensoñaciones. Estos autores sugieren, plausiblemente, que las autoexhortaciones y los recuerdos, posibles gracias al lenguaje, son realmente esenciales para el mantenimiento del tipo de proyectos a largo plazo en los que sólo los humanos podemos implicarnos (a menos que, como el castor, tengamos un constructor incorporado para completar un proyecto concreto a largo plazo). (Para más discusiones sobre estos temas, véanse Clark y Karmiloff-Smith 1994 Y Dennett 1994c).

Esto me lleva al escalón final en mi torre de generación y prueba. Hay una incorporación más de esta admirable idea, y es aquella que proporciona a nuestras mentes el mayor de los poderes: una vez que tenemos lenguaje —una caja repleta de herramientas mentales— podemos utilizar estas herramientas en la estructura que *genera* y que somete a *prueba*, de manera premeditada y con visión de futuro, conocida como *ciencia*. Todas las otras variedades de generación y prueba son un «quiero y no puedo».

El soliloquio que acompaña los errores cometidos por las más bajas criaturas skinnerianas puede ser, «Bien, ¡no debo hacer *esto* otra vez!», y la más dura lección que debe aprender cualquier agente, aparentemente, es cómo aprender de sus propios errores. Para aprender de ellos, hay que estar dispuesto a contemplarlos y esto no es un asunto menor. La vida se mueve impetuosamente y, a menos que hayamos desarrollado estrategias positivas para recordar un recorrido, la tarea conocida en la inteligencia artificial como *asignación de crédito* (también conocida, naturalmente, como «asignación de culpabilidad») es insoluble. El descubrimiento de la fotografía de alta velocidad fue un extraordinario avance tecnológico de la ciencia, ya que

permitió a los seres humanos, por primera vez, examinar complicados fenómenos temporales no en tiempo real sino *en su propio tiempo*, sin prisas, con un análisis retroactivo y metódico de los caminos que han originado aquellos complicados sucesos. En este caso, un avance tecnológico nos aporta una enorme potenciación del poder cognitivo. La llegada del lenguaje fue un *boom* exactamente paralelo al de la fotografía de alta velocidad, para los seres humanos, una tecnología que ha creado una clase totalmente nueva de objetos para contemplar, verbalmente incorporados, que pueden revisarse en cualquier orden y a cualquier ritmo. Y esto abre una nueva dimensión de automejoramiento; todo lo que uno tenía que hacer era aprender a saborear sus propios errores.

Sin embargo, la ciencia no es justamente un asunto de cometer errores sino de cometer errores en público. Cometer errores para que los vean todos, con la esperanza de conseguir ayuda de otros con las correcciones. Nicholas Humphrey, David Premack [1986] y otros sostienen que los chimpancés son psicólogos naturales, lo que yo llamaría sistemas intencionales de segundo orden, capaces de adoptar la posición intencional hacia otras cosas. Esto no es sorprendente si nuestro propio e innato equipamiento incluye una teoría del módulo mental, como Leslie, Baron-Cohen y otros han afirmado, que quizás es parte de la dotación que los chimpancés y nosotros hemos heredado de un ancestro común. Pero aunque los chimpancés están, como nosotros, equipados de forma innata como psicólogos naturales, les falta una característica crucial compartida por todos los humanos que son psicólogos naturales, en sus variedades de psicología popular y profesional: nunca consiguen comparar notas. Nunca disputan sobre atribuciones ni desean conocer los fundamentos de las conclusiones de otros. No nos extraña que su compresión sea tan limitada. La nuestra lo sería si tuviéramos que generarlo todo a partir de nosotros mismos.

Permítaseme hacer un resumen de los resultados de este recorrido algo rápido. Nuestros cerebros humanos, y sólo los cerebros humanos, han sido equipados con hábitos y métodos, herramientas de la mente e información derivada de millones de otros cerebros que no son ancestros de nuestros cerebros. Todo esto, amplificado por el uso premeditado y previsor de la *generación* y la *prueba* en el contexto de la ciencia, ha elevado nuestras mentes a un plano diferente de las mentes de nuestros parientes más cercanos entre los animales. Este específico proceso de potenciación de nuestra especie ha llegado a ser tan rápido y poderoso que una simple generación de sus mejoras en el diseño puede ahora empequeñecer los esfuerzos de I+D de

millones de años de evolución mediante la selección natural. Comparar anatómicamente nuestros cerebros con los cerebros de los chimpancés —o con los cerebros de los delfines u otros cerebros no humanos— casi no viene al caso, debido a que nuestros cerebros están, en efecto, unidos por un sistema cognitivo singular que empequeñece a los demás. Están unidos por una innovación que ha invadido nuestros cerebros y no los otros: el lenguaje. No estoy afirmando estúpidamente que todos nuestros cerebros están entrelazados por el lenguaje hasta formar una gigantesca mente, pensando pensamientos transnacionales, sino que cada cerebro humano, individualmente considerado y gracias a sus lazos comunicativos, es beneficiario de las labores cognitivas de los otros, de un modo que nos otorga poderes sin precedentes.

Los desnudos cerebros animales no son comparables, en modo alguno, con los cerebros fuertemente armados y pertrechados que portamos en nuestras cabezas. Este hecho invierte la carga de la prueba en lo que de otro modo sería un argumento apremiante: la tesis, defendida primeramente por el lingüista Noam Chomsky [1975] y más recientemente por los filósofos Jerry Fodor [1983] y Colin McGinn [1991], que sostiene que nuestras mentes, como las de todas las otras especies, deben sufrir un «cierre cognitivo» con respecto a la investigación de algunos temas. Las arañas no pueden contemplar el concepto de pesca; los pájaros (algunos de los cuales son excelentes pescadores) no pueden pensar en la democracia. Lo que es inaccesible al perro o al delfín puede ser fácilmente captado por el chimpancé, pero, a su vez, el chimpancé estaría cognitivamente cerrado a algunos dominios acerca de los cuales nosotros, los humanos, podemos pensar. Chomsky y compañía plantean una cuestión retórica: ¿Qué nos hace pensar que somos diferentes? ¿Estamos confinados en límites estrictos en lo que se refiere a lo que el *Homo sapiens* puede concebir?

De acuerdo con Chomsky, todos los temas correspondientes a las perplejidades humanas pueden agruparse en «problemas» cuando pueden ser resueltos y en «misterios» cuando no se pueden resolver. El problema del libre albedrío, opina Chomsky, es uno de esos misterios[134]. El problema de la conciencia, según Fodor, es otro y McGinn está de acuerdo. Como autor de libros (Dennett 1984, 1991a) que sostienen que se pueden explicar cada uno de estos impenetrables misterios, puede esperarse que yo no esté de acuerdo, pero éste no es el lugar para proseguir con estos temas. Dado que ninguno de los dos, ni Chomsky ni Fodor, piensan que puedan explicar el libre albedrío o la consciencia, la tesis de que esto es humanamente imposible resulta

doctrinalmente conveniente para ellos, quizá, pero también entra en considerable tensión con otras de sus afirmaciones. Con otro talante, ambos han aclamado (acertadamente) la capacidad del cerebro humano para «analizar» y, en consecuencia, presumiblemente entender, la oficial infinitud de oraciones gramaticales de un lenguaje natural como el inglés. Si podemos comprender, en principio, todas las frases, ¿no podríamos comprender los ordenados grupos de sentencias que mejor expresan la solución de los problemas del libre albedrío y de la consciencia? Después de todo, uno de los volúmenes de la Biblioteca de Babel es —debe ser— la mejor declaración en menos de quinientas páginas de frases gramaticales cortas en inglés acerca de la solución del problema del libre albedrío, y otro, el óptimo trabajo en inglés sobre la consciencia[135]. Yo no tengo duda de que ninguno de mis libros es uno de aquellos, pero así es la vida. Yo no puedo creer que Chomsky y Fodor declarasen que cualquiera de aquellos libros (o el trillón de subcampeones) era incomprensible para un lector normal en inglés[136]. Así que quizás ellos piensen que los misterios del libre albedrío y de la consciencia son tan profundos que ningún libro —por grande que sea, en cualquier lenguaje— podría explicárselos a un ser inteligente. Pero *esta* afirmación no tiene a su favor evidencia alguna que se derive de cualquier consideración biológica. ¡Debe de haber caído del cielo!

Consideremos con más detalle el argumento del «cierre cognitivo». «Lo que está cerrado a la mente de una rata puede estar abierto para la mente de un mono, y lo que está abierto para nosotros puede estar cerrado para un mono» (McGinn 1991:3). Los monos, por ejemplo, no pueden captar el concepto de un electrón, nos recuerda McGinn, pero yo creo que no debemos sentirnos impresionados por este ejemplo, sólo porque el mono no pueda comprender las respuestas respecto a los electrones, sino porque el mono no puede comprender las preguntas (Dennett 1991d). El mono *no está desconcertado* en absoluto. Nosotros definitivamente comprendemos las preguntas acerca del libre albedrío y la consciencia lo suficientemente bien para reconocer que estamos desconcertados (si lo estamos), así que hasta que Chomsky y Fodor y McGinn puedan aportarnos casos claros de animales (o gentes) que pueden desconcertarse ante preguntas cuyas verdaderas respuestas no puedan no desconcertarlos, no nos han proporcionado evidencia acerca de la realidad, y ni siquiera la posibilidad, del «cierre cognitivo» en los seres humanos[137].

Su argumento se presenta como un argumento biológico, naturalista, al recordarnos nuestros parentescos con los otros animales y aconsejarnos no

caer en la vieja trampa de pensar que nosotros, «almas» humanas, somos «como ángeles», con nuestras mentes «infinitas». Pero éste es de hecho un argumento pseudobiológico, un argumento que, al ignorar los reales detalles biológicos, nos dirige equivocadamente fuera del caso, al sacar una especie —nuestra especie— totalmente fuera de la escala de inteligencia que clasifica a los cerdos por encima de los lagartos y a la hormiga por encima de la ostra. Es cierto que, en principio, no podemos eliminar la posibilidad de que nuestras mentes estén cerradas cognitivamente a uno u otro dominio. De hecho, como hemos visto con más detalle en el capítulo 15, podemos estar seguros de que hay territorios de conocimientos, importantes y fascinantes, en los que nuestra especie, en su real finitud, nunca entrará, no porque golpeemos nuestra cabeza contra alguna muralla de absoluta incomprensión sino debido a que la muerte por calor del universo nos liquidará antes de que podamos conseguirlo. Sin embargo, ésta no es una limitación debida a la debilidad de nuestros cerebros animales, ni un dictado del «naturalismo». Por el contrario, la aplicación apropiada del pensamiento darwiniano sugiere que, si sobrevivimos a las actuales crisis del medio ambiente inducidas por nosotros mismos, nuestra capacidad de comprender continuará creciendo a través de incrementos que ahora son incomprensibles para nosotros.

¿Por qué no gustará esta conclusión a Chomsky, a Fodor ni a McGinn? Esta conclusión confía en las mentes humanas —y solamente en las mentes humanas— en un indefinido dominio en expansión a través de los rompecabezas y los problemas del universo, sin límites a la vista. ¿Qué puede ser más admirable que esto? El problema es, sospecho, que ellos consideran que los *medios* son insatisfactorios; si el poder de la mente se debiera a las grúas, no a los ganchos celestes, ellos estarían dispuestos a aceptar el misterio. En cualquier caso, esta actitud sale a la superficie en estas controversias y Chomsky es una fuente primaria de autoridad a este respecto.

2. *Chomsky contra Darwin: cuatro episodios*

> Para Chomsky, se podría pensar, todo sería ganancia si asentase su controvertida teoría acerca de un órgano del lenguaje en los firmes fundamentos de la teoría de la evolución, y en algunos de sus escritos ha mencionado, de un modo indirecto, esta conexión. Pero, en general, se muestra escéptico.
>
> Steven Pinker, *El instinto del lenguaje*

> En el caso de sistemas tales como el lenguaje o las alas ni siquiera es fácil imaginar un recorrido de la selección que hubiera podido originarlos.
>
> Noam Chomsky,
> *Language and Problems of Knowledge: the Managua Lectures*

> Aún existe un considerable abismo de comunicación entre las ciencias cognitivas que entraron en el campo desde la IA, o desde el estudio de la conducta ante la solución de problemas y la formación de conceptos, por un lado, y por otro, aquellos que entraron desde la preocupación por el lenguaje… Cuando se subraya enfáticamente el carácter único del proceso del lenguaje como facultad humana, como ha hecho Chomsky… el abismo se agranda.
>
> Herbert Simon y Craig Kaplan,
> *«Foundations on Cognitive Science»*

El 11 de septiembre de 1956, en el Massachusetts Institute of Technology (MIT) se presentaron tres ponencias en una reunión del Institute of Radio Engineers. Una, de Alien Newell y Herbert Simon [1956], titulada «The Logic Theory Machine» (‘La máquina de la teoría lógica’) mostraba, por vez primera, cómo un ordenador podía probar teoremas no triviales de la lógica. La «máquina» de Newell y Simon era la madre (o la abuela) de su General Problem Solver (‘Solucionador de problemas generales’) (Newell y Simon 1963) y del prototipo del ordenador del lenguaje Lisp que es a la inteligencia artificial lo que aproximadamente el código del ADN es a la genética. La máquina de la teoría lógica es una digna rival del programa de damas de Art Samuel. La otra ponencia se titulaba «The Magical Number Seven, Plus or Minus Two» (‘El mágico número siete, más o menos dos’) y su autor era el psicólogo George A. Miller, y se convirtió en uno de los trabajos clásicos que inauguraron el campo de la psicología cognitiva (Miller 1956). La tercera comunicación fue presentada por un joven de veintisiete años, *junior Fellow* en Harvard, llamado Noam Chomsky y se titulaba «Three Models for the Description of Language» (1956). Las coronaciones retrospectivas son siempre un poco arbitrarias, como hemos comentado varias veces, pero la conferencia de Chomsky en el Institute of Radio Engineers es un acontecimiento tan importante que marcó el nacimiento de la lingüística moderna. Tres importantes nuevas disciplinas científicas nacieron en la misma sala de conferencias y el mismo día; me pregunto cuántos de los asistentes tuvieron la sensación de estar participando en un acontecimiento histórico de tales proporciones. George Miller la tuvo, como nos cuenta en su

relato (1979) de dicha reunión. La visión retrospectiva de la ocasión del propio Herbert Simon ha cambiado a lo largo de los años. En su libro de 1969, Simon se ocupó de esta ocasión extraordinaria y dijo: «Los dos cuerpos teóricos (lingüística e inteligencia artificial) han mantenido relaciones cordiales desde el principio. Y desde luego, permanecen conceptualmente en la misma visión del cerebro humano» (1969:47). ¡Ojalá fuera verdad! En el año 1989, él mismo vería cómo se había agrandado el abismo.

Pocos científicos son grandes científicos, y pocos grandes científicos llegan a encontrar un campo totalmente nuevo, aunque hay unos cuantos. Charles Darwin es uno; Noam Chomsky es otro. Del mismo modo que había biología antes de Darwin —historia natural y fisiología y taxonomía—, todas ellas unidas por Darwin en lo que conocemos hoy como biología, también había lingüística antes de Chomsky. El campo científico contemporáneo de la lingüística con sus subdisciplinas de fonología, sintaxis, semántica y pragmática, sus belicosas escuelas y sus renegados disidentes (la lingüística computacional en la IA, por ejemplo), sus subdisciplinas de psicolingüística y neurolingüística, creció a partir de varias tradiciones académicas hasta llegar retrospectivamente a los pioneros investigadores y teóricos del lenguaje, desde los hermanos Grimm hasta Ferdinand de Saussure y Roman Jakobson, aunque todo esto quedó unificado en una rica e interrelacionada familia de investigaciones científicas por los avances teóricos primariamente propuestos por un pionero, Noam Chomsky. Su escueto libro de 1957 titulado *Syntactic Structures* ('Estructuras sintácticas') fue una aplicación a los lenguajes naturales, como el inglés, de los resultados de una ambiciosa investigación teórica que él había emprendido en otra región del espacio de diseño: el espacio lógico de todos los posibles algoritmos para generar y reconocer las sentencias de todos los lenguajes posibles. El trabajo de Chomsky siguió de cerca el camino de las investigaciones puramente lógicas de Turing sobre los poderes de lo que ahora llamamos ordenadores. Chomsky definió finalmente una escala ascendente de tipos de gramáticas o tipos de lenguaje —la jerarquía de Chomsky, con la cual todos los estudiantes de la teoría de la computación inician sus estudios— y demostró cómo estas gramáticas eran interdefinibles en una escala ascendente de tipos de autómatas u ordenadores, desde las máquinas de estados finitos a los autómatas *push-down*, las máquinas ligadas a tiempo lineal y a las máquinas de Turing.

Puedo recordar claramente la onda de choque que se extendió a través de la filosofía cuando el primer ensayo de Chomsky llegó a nuestras manos, pocos años después. En 1960, mi primer año en Harvard, pregunté al profesor

Quine quiénes eran los críticos de sus puntos de vista que yo debía leer (en aquel tiempo me consideraba un anti quiniano de feroz convicción y con vistas a mi tesis como *senior*; iniciaba el desarrollo de argumentos que lo atacaran. ¡Debía conocer a todo aquel que hubiera argumentado contra Quine!). Quine inmediatamente me sugirió que debía leer el trabajo de Noam Chomsky, un autor del que pocos filósofos habían oído hablar en aquel tiempo, aunque su fama pronto nos envolvería. Los filósofos del lenguaje se dividieron entre los que le amaban y los que le odiaban. Los primeros pronto estuvimos hasta las cejas de transformaciones, árboles, estructuras profundas y todos los arcanos de un nuevo formalismo. Muchos de aquellos que le odiaban lo condenaban como un terrible representante del *cientificismo* filisteo, decían que se trataba de un estrepitoso asalto, por vándalos tecnócratas, de las bellas, inanalizables e informalizables sutilezas del lenguaje. Esta actitud hostil fue potenciada en los departamentos de lenguas extranjeras de la mayor parte de las universidades. Chomsky podía ser un profesor de lingüística en el MIT, y la lingüística podía ser considerada en esa institución como una de las humanidades, pero el trabajo de Chomsky era ciencia, y la ciencia era el enemigo, como sabe todo humanista con carnet.

> Dulce es la enseñanza que la Naturaleza nos proporciona;
> nuestro entrometido intelecto
> desfigura las bellas formas de las cosas:
> asesinamos cuando diseccionamos.

La romántica visión que tenía Wordsworth del científico como asesino de la belleza parece encarnar perfectamente en Noam Chomsky, teórico de los autómatas e ingeniero de radio, pero no deja de ser una gran ironía que él, que ha sido el defensor de una acritud a favor de la ciencia, parezca ofrecer la salvación a los humanistas. Como hemos visto en la sección anterior, Chomsky ha argumentado que la ciencia tiene límites y que, en particular, cuando se aplica al entendimiento «se pilla los dedos». Discernir cómo se ha configurado este curioso hecho ha sido largo y difícil, aun para aquellos capaces de manejar las complejidades técnicas y las controversias de la lingüística contemporánea, aunque durante largo tiempo nos ha maravillado. La famosa revisión de Chomsky [1959] criticando severamente el libro de B. F. Skinner titulado *Verbal Behavior*, publicado en 1957, fue uno de los documentos fundadores de la ciencia cognitiva. Al mismo tiempo, Chomsky se había mantenido firmemente hostil a la inteligencia artificial, y fue lo bastante atrevido como para titular uno de sus libros más importantes, publicado en 1966, *Cartesian Linguistics*, casi como si pensara que el

dualismo antimaterialista de Descartes volvería a ponerse de moda. En todo caso, ¿de qué lado está Chomsky? En cualquier caso, no del lado de Darwin. Si los que temen a Darwin desean un defensor que esté profundamente enraizado, y con gran influencia, en la ciencia, no podrán encontrar a otro mejor que Chomsky.

Me costó tiempo caer en la cuenta. En marzo de 1978 fui el anfitrión en Tufts de un excelente debate, organizado por la Society of Philosophy and Psychology[138]. Lo que era nominalmente un panel para discutir sobre los fundamentos y las perspectivas de la inteligencia artificial, se convirtió en un combate de lucha retórica entre cuatro ideólogos que eran pesos pesados de la ideología: Noam Chomsky y Jerry Fodor atacando la inteligencia artificial y Roger Schank y Terry Vinograd defendiéndola. Por aquel tiempo, Schank estaba trabajando en un programa para la comprensión del lenguaje natural y los críticos localizaron sus ataques en su esquema para representar (en un ordenador) la enormemente confusa colección de trivialidades que todos conocemos y en la que, de algún modo, confiamos cuando desciframos actos ordinarios del lenguaje, cuando son alusivos y truncados. Chomsky y Fodor acumularon desdén en esta empresa, aunque los fundamentos de sus ataques se desplazaron gradualmente en el curso del combate, ya que Schank no es ningún perezoso cuando se trata de colocar cebos en el anzuelo, y defendió tenazmente su proyecto de investigación. Los ataques de Chomsky y Fodor comenzaron con una condena del error conceptual, muy directa y basada en «primeros principios» —Schank iba de un lado a otro, como un loco—, aunque terminó con una llamativa concesión de Chomsky: podría suceder, como pensaba Schank, que la capacidad humana para comprender una conversación (o, de modo más general, para pensar) tuviera que explicarse en términos de interacción de cientos o miles de artilugios de pacotilla, pero esto sería una vergüenza, ya que probaría que la psicología, al fin, no era «interesante». En la mente de Chomsky sólo había dos posibilidades interesantes: la psicología podría llegar a ser «como la física», con sus regularidades explicables como consecuencias de unas pocas leyes, profundas, elegantes e inexorables; o bien la psicología podría estar absolutamente desprovista de leyes, en cuyo caso la única manera de estudiar o exponer la psicología sería la del novelista (y Chomsky tenía más preferencias por Jane Austen que por Roger Schank, si llegara el caso).

Siguió un fuerte debate entre los panelistas y la audiencia, rematado por una observación de Marvin Minsky, colega de Chomsky en el MIT: «Creo que sólo un profesor de humanidades en el MIT puede olvidarse de una

“tercera e interesante posibilidad”: que la psicología sea como la ingeniería». Minsky había puesto el dedo en la llaga. En la perspectiva de un planteamiento de la mente a partir de la ingeniería, hay algo que resulta profundamente repugnante para un cierto tipo de humanistas, y esto no tiene nada que ver con una aversión al materialismo o a la ciencia. El mismo Chomsky es un científico y presumiblemente un materialista (¡su lingüística *cartesiana* no está muy lejos de esta posición!), pero no tendría tratos con la ingeniería. Sería, de alguna manera, algo por debajo de la dignidad del intelecto que ésta fuera un artefacto o una colección de artefactos. Es mejor que la mente vuelva a ser un misterio impenetrable, un *sanctasanctórum* para el caos, que llegase a convertirse en un tipo de entidad que pudiera proporcionarnos sus secretos a través de los análisis de la ingeniería.

Aunque quedé impactado, en aquel tiempo, por la observación de Minsky sobre Chomsky, no comprendí el mensaje en su totalidad. En 1980, Chomsky publicó un artículo titulado «Rules and Representations». (‘Reglas y representaciones’), específicamente abierto a discusión en la revista *Behavioral and Brain Sciences* (BSS), y yo fui uno de sus comentadores. El contencioso era, y sigue siendo, la insistencia de Chomsky en que la competencia del lenguaje era, en su mayor parte, innata, no algo que un niño pudiera *aprender*. Según Chomsky, la estructura del lenguaje está firmemente fijada en forma de reglas innatas específicas, y todos los niños la tienen en unos pocos «interruptores» más bien periféricos que lo convierten en un inglés-parlante en lugar de en un chino-parlante. Chomsky dice que el niño *no* es una especie de aprendiz para todo —un solucionador de todo tipo de problemas, como dirían Newell y Simon— obligado a resolver de qué lenguaje se trata y a aprenderlo. Más bien, el niño se encuentra equipado de modo innato para hablar y comprender un lenguaje y, sencillamente, ha de excluir ciertas posibilidades, muy limitadas, e incluir otras. Esta es la razón por la que los niños aprenden a hablar sin esfuerzo, incluso los calificados como «lentos». Realmente no están aprendiendo, no más que los pájaros aprenden sus plumas. El lenguaje, y las plumas, se desarrollan precisamente en especies que están ordenadas para tenerlos y quedan fuera de los límites de las especies a las que la falta este equipamiento innato. Unos pocos disparadores del desarrollo ponen en marcha el proceso de adquisición del lenguaje y unas cuantas condiciones ambientales aportan subsiguientemente algunas podas o conformaciones de menor cuantía, dentro de la lengua materna con la que el niño se encuentra.

Esta tesis ha encontrado enorme resistencia, pero ahora estamos seguros de que la verdad se encuentra más cerca del extremo de la mesa donde se encuentra Chomsky que en el extremo de sus oponentes (para más detalles, véanse las defensas de la posición de Chomsky, en Jackendoff 1993 y Pinker 1994). ¿Por qué esta resistencia? En mi comentario publicado en la revista *BBS* —que presenté como una observación constructiva y no como una objeción—, yo señalaba que había una razón para ofrecer resistencia a Chomsky que era perfectamente razonable, aunque sólo se tratara de una razonable *esperanza*. Del mismo modo que en la resistencia de los biólogos a la hipótesis del astrónomo Hoyle —la hipótesis de que la vida no comenzó en la Tierra, sino en algún otro lugar desde donde emigró hasta aquí—, la resistencia de los psicólogos al reto de Chomsky tenía una explicación: si Chomsky estaba en lo cierto, esto significaba que el fenómeno del lenguaje y la adquisición del lenguaje resultaría mucho más difícil de investigar. En lugar de encontrarnos con el proceso de aprendizaje en marcha, ante nuestros ojos, en niños concretos, donde pudiéramos estudiarlo y manipularlo, tendríamos que «traspasarle la responsabilidad a la biología», y esperar a que los biólogos nos explicaran cómo «aprendió» nuestra especie a tener incorporadas, en el momento de nacer, las competencias del lenguaje. Este es un programa de investigación mucho menos manejable. En el caso de la hipótesis de Hoyle, se podrían imaginar

argumentos que fijasen una máxima velocidad de mutación y selección y demostrando que había habido tiempo suficiente en la Tierra para que la *totalidad* del proceso hubiera ocurrido localmente.

Los argumentos de Chomsky, desde la pobreza del estímulo y de la velocidad de la adquisición del lenguaje, son análogos; pretenden demostrar que debe haber habido en el niño grandes regalos de diseño, si tuviésemos que explicar el rapidísimo desarrollo de la competencia madura. Y aunque pudiéramos consolarnos con la suposición de que podremos ser capaces algún día de confirmar la presencia de estructuras innatas en el cerebro, mediante el examen directo del sistema nervioso (algo así como encontrar fósiles de nuestros ancestros extraterrestres) tendremos que aceptar la descorazonadora conclusión de que la mayor parte de lo que hemos esperado de la teoría del aprendizaje, considerada en su forma más general como el intento de explicar la transición de la absoluta ignorancia al conocimiento, no es, en absoluto, la provincia de la psicología, sino de la biología evolucionista más especulativa (Dennett 1980).

Para mi sorpresa, Chomsky no comprendió el verdadero sentido de mi comentario. Mientras que él mismo había ofrecido reflexiones sobre lo que haría «interesante» a la psicología, era incapaz de ver hasta qué punto podía ser descorazonador para los psicólogos el descubrimiento de que la responsabilidad pudiera pasar a la biología. ¡Años más tarde, comprendí que la razón por la que Chomsky no se dio cuenta de cuál era el sentido de mi comentario era que mientras él insistía en que el «órgano del lenguaje» era

innato, esto *no* significaba para él que fuera un producto de la selección natural! O, al menos, no lo era de una forma que permitiera a los biólogos *recoger* la responsabilidad y analizar la vía mediante la cual el medio ambiente de nuestros ancestros había conformado el diseño del órgano del lenguaje a través de los eones. El órgano del lenguaje, pensaba Chomsky, no era una adaptación sino… un misterio, o un monstruo prometedor. Era algo que quizá sería clarificado algún día por la física, pero no por la biología.

Puede ser que en algún remoto período haya ocurrido una mutación que dio origen a la propiedad de un discreto infinito, quizá por razones que tienen que ver con la biología de las células, para ser explicada en términos de propiedades de mecanismos físicos, ahora desconocidos… Muy posiblemente otros aspectos de este desarrollo evolutivo reflejan de nuevo la operación de leyes físicas aplicándose a un cerebro de un cierto grado de complejidad (1988:170).

¿Cómo pudo ser esto? Muchos lingüistas y biólogos han perseguido las huellas del problema de la evolución del lenguaje, utilizando los mismos métodos que han funcionado bien en otros rompecabezas evolutivos y consiguiendo resultados o, al menos, lo que parecen ser resultados. Por ejemplo, en el extremo más empírico del espectro, los trabajos de los neuroanatómicos y los psicolingüistas han demostrado que nuestros cerebros tienen características que están ausentes en los cerebros de nuestros parientes más cercanos supervivientes, características que desempeñan funciones cruciales en la percepción y en la producción del lenguaje. Hay una amplia diversidad de opiniones acerca de cuándo en los últimos seis millones de años, más o menos, nuestro linaje adquirió estas características, en qué orden y por qué, pero estos desacuerdos son tan tratables en ulteriores investigaciones —ni mejores ni peores— como las discrepancias acerca del *archaeopteryx*, por ejemplo. En un planteamiento puramente teórico, y extendiendo la red más ampliamente, las condiciones para la evolución de los sistemas de comunicación en general han sido deducidas (por ejemplo, en Krebs y Dawkins 1984, Zahavi 1987) y sus implicaciones exploradas en modelos de simulación y en experimentos empíricos.

Ya vimos en el capítulo 7 algunas de las ingeniosas especulaciones y modelos utilizados frente al problema de cómo la vida, sin ayuda ajena, empezó a existir, y hay una similar generosidad de ideas inteligentes acerca de cómo el lenguaje se puso en marcha. No hay duda de que el origen del lenguaje es teóricamente un problema mucho más fácil que el origen de la vida; tenemos un rico catálogo de materiales, no tan crudos, con los que construir una respuesta. Puede que nunca podamos confirmar los detalles, pero, si esto es así, no será un misterio, sino solamente un *bit* de irreparable

ignorancia. Algunos científicos especialmente abstemios pueden ser reticentes a dedicar tiempo y atención a tales ejercicios de largos vuelos en la especulación deductiva, pero ésta no parece ser la posición de Chomsky. Sus reservas se dirigen no a la probabilidad de éxito, sino al objetivo de la empresa.

Es perfectamente seguro atribuir este desarrollo de estructuras innatas del lenguaje a la «selección natural», con tal de que caigamos en la cuenta de que no hay sustancia para esta afirmación, que no llega a ser más que una creencia, y que hay alguna explicación naturalista para este fenómeno (Chomsky 1972:97).

Desde hace tiempo ha habido signos del agnosticismo de Chomsky —e incluso antagonismo— hacia el darwinismo, aunque muchos lo encontramos difícil de interpretar. Para algunos, Chomsky aparece como un «neocreacionista», aunque esto no parece muy plausible, especialmente desde que tiene el apoyo de Stephen Jay Gould. ¿Recordáis la apelación del lingüista Jay Keyser al término *spandrel* de Gould, para describir cómo el lenguaje llegó a ser? Es probable que Keyser consiguiera esta terminología de su colega Chomsky, quien la obtuvo de Gould, quien, como agradecimiento, ha apoyado ansiosamente el punto de vista de Chomsky de que el lenguaje no evolucionó realmente, sino que, más bien, llegó por sorpresa, como un regalo inexplicable, a lo mejor un producto secundario del agrandamiento del cerebro humano.

Sí, el cerebro se hizo grande por selección natural. Pero como resultado de este tamaño, y de la densidad nerviosa y de las numerosísimas conexiones que esto implicaba, los cerebros humanos podían llevar a cabo un inmenso espectro de funciones no del todo relacionadas con las razones originales de su aumento de volumen. El cerebro no se hizo grande para que pudiera leer, escribir o hacer operaciones aritméticas o establecer el ciclo de las estaciones; porque la cultura humana, como sabemos, depende de habilidades de este tipo… Los universales del lenguaje son tan diferentes de cualquier otra cosa de la naturaleza, y tan singulares en su estructura, que se originan como una consecuencia colateral del incremento de la capacidad cerebral, más que como un simple avance en continuidad desde los gruñidos y gestos ancestrales, como se había indicado. (Este argumento acerca del lenguaje no es original mío, aunque me adhiero totalmente a él; esta línea de razonamiento se sigue directamente como lectura evolucionista de la teoría de Noam Chomsky de la gramática universal.) (Gould 1989b:14).

Gould subraya que el crecimiento del cerebro puede no haberse debido inicialmente a la selección para el lenguaje (o incluso para un aumento de la inteligencia) y que el lenguaje humano puede no haberse desarrollado «como un simple avance en continuidad desde los gruñidos ancestrales», aunque no se sigue de estas suposiciones (las cuales podemos aceptar sin pruebas para esclarecer el tema de esta discusión) que el órgano del lenguaje no sea una adaptación. Es, permítaseme, una exadaptación, pero las exadaptaciones son

adaptaciones. ¡Dejemos que el notable crecimiento del cerebro del homínido sea una pechina en el sentido que Gould o Keyser desean, y *aún* el órgano del lenguaje seguirá siendo una adaptación igual que el ala del pájaro! No importa con qué rapidez haya ocurrido la puntuación que empujara abruptamente a nuestros ancestros a la derecha en el espacio de diseño, para que haya un desarrollo gradual del diseño bajo la presión de la selección natural, a menos que se trate de un milagro o de un monstruo prometedor. En resumen, mientras Gould proclama la teoría de Chomsky de la gramática universal como un baluarte contra una explicación adaptacionista del lenguaje, y Chomsky, en recompensa, apoya la postura antiadaptacionista de Gould como una excusa de autoridad para rechazar la obvia obligación de buscar afanosamente una explicación evolucionista del establecimiento innato de la gramática universal, estas dos autoridades se han sostenido una a otra sobre un abismo.

En diciembre de 1989, Steven Pinker, psicolingüista del MIT, y su alumno graduado Paul Bloom presentaron una ponencia titulada «Natural Language and Natural Selection» (‘Lenguaje Natural y Selección Natural’) en el Coloquio de Ciencia Cognitiva celebrado en el MIT. Este trabajo, que apareció posteriormente como artículo para discusión en la revista BBS y lanzó el guante:

> Mucha gente ha argumentado que la evolución de la facultad humana del lenguaje no puede explicarse por la selección natural de Darwin. Chomsky y Gould han sugerido que el lenguaje pudo haber evolucionado como un producto secundario de selección a partir de otras habilidades, o como una consecuencia de leyes todavía desconocidas de crecimiento y forma… Nosotros concluimos que hay muchas razones para creer que una especialización en la gramática evolucionó mediante un proceso neodarwiniano convencional (Pinker y Bloom 1990:707).

«En cierto sentido», dicen Pinker y Bloom, «nuestro objetivo es increíblemente aburrido. Todo lo que argumentamos es que el lenguaje no es diferente de otras complejas habilidades tales como la detección mediante ultrasonidos o la visión estereoscópica, y que la única vía para explicar el origen de tales habilidades es a través de la teoría de la selección natural» (1990:708). Llegaron a esta «increíblemente aburrida» conclusión a través de una paciente evaluación de varios análisis de fenómenos muy diversos que muestran, más allá de toda duda razonable —¡sorpresa, sorpresa!—, que el órgano del lenguaje debe haber evolucionado muchas de sus más interesantes propiedades como adaptaciones, justamente como esperaban algunos neodarwinistas. La respuesta de la audiencia en el MIT fue, sin embargo, de todo menos aburrida. Estaba previsto que Chomsky y Gould respondieran, así

que había una multitud de pie en la sala[139]. Me escandalizó el nivel de hostilidad e ignorancia acerca de la evolución que, en aquella ocasión, expresaron desenfadadamente eminentes científicos cognitivos. (De hecho, fue la reflexión sobre lo sucedido en dicha reunión lo que me persuadió de que no podía pasar más tiempo sin escribir, este libro). Hasta donde yo sé, no existe una transcripción de aquella reunión (los comentarios en la revista BBS incluyen algunos de los temas planteados en la reunión) pero se puede recuperar algo del tono de la reunión revisando la lista de Pinker (comunicación personal) con las diez más asombrosas objeciones que él y Bloom contestaron satisfactoriamente, desde que comenzaron a circular copias de su trabajo. La mayoría de sus versiones, si la memoria no me falla, se expresaron en la reunión del MIT;

1) La visión de los colores no tiene función alguna; podríamos separar las manzanas rojas de las manzanas verdes utilizando señales de intensidad.

2) El lenguaje no ha sido diseñado en modo alguno para la comunicación: no es como un reloj, sino como un artilugio de Rube Goldberg con una varilla en el medio que podemos utilizar como un reloj de sol.

3) Cualquier argumento a favor de que el lenguaje es funcional puede formularse con igual plausibilidad y fuerza cuando se aplica a escribir en la arena.

4) La explicación de la estructura de la célula debe ser aportada por los físicos, no por la evolución.

5) Tener ojo exige el mismo tipo de explicación que tener masa, debido a que así como el ojo nos permite ver, la masa nos previene de flotar en el espacio.

6) ¿No ha refutado a Darwin ese disparate acerca de las alas de los insectos?

7) El lenguaje no puede ser útil; conduce a la guerra.

8) La selección natural es irrelevante, debido a que ahora tenemos la teoría del caos.

9) El lenguaje no pudo haber evolucionado a través de la presión de la evolución para la comunicación, ya que podemos preguntar a la gente cómo se siente sin realmente desear saberlo.

10) Todo el mundo está de acuerdo en que la selección natural desempeña algún papel en el origen de la mente pero esto no puede explicar todos los aspectos, así que no hay nada más que decir.

¿Son Gould y Chomsky responsables de las extrañas convicciones de algunos de sus partidarios? Esta cuestión no tiene una respuesta simple. Más

de la mitad de los puntos en la lista de Pinker tiene claros antecedentes en las afirmaciones hechas por Gould (números 2, 6 y 9 en particular) y Chomsky (números 4, 5 y 10 en particular). Aquellos que hacen estas afirmaciones (incluyendo las otras en la lista) las presentan característicamente bajo la autoridad de Gould y Chomsky (véase, por ejemplo, Otero 1990). Como dicen Pinker y Bloom, «Noam Chomsky, el más grande lingüista mundial, y Stephen Jay Gould, el teórico evolucionista mundialmente más conocido, han sugerido en repetidas ocasiones que el lenguaje no puede ser el producto de la selección natural» (1990:708). Además —dos importantes perros que no han ladrado— tengo todavía que ver a Gould o a Chomsky intentando corregir a esos aulladores cuando se levanten en el fragor de la batalla. (Como veremos, esta es la debilidad de todos; yo lamento que la mentalidad de sitio entre los sociobiólogos haya conducido a insignes malos razonamientos por parte de los miembros de su propio equipo).

Uno de los partidarios más entusiastas de Darwin era Herbert Spencer, acuñador de la frase «supervivencia del mejor adaptado» y un importante esclarecedor de algunas de las mejores ideas de Darwin, pero también el padre del darwinismo social, una odiosa y errónea aplicación del pensamiento darwinista en defensa de doctrinas políticas que varían desde lo insensible a lo nefando**[140]**. ¿Era Darwin responsable del mal uso que hizo Spencer de sus puntos de vista? Las opiniones difieren sobre este punto. Por mi parte, excuso a Darwin de la tarea verdaderamente heroica de castigar a su defensor en público, aun cuando lamento que no fuera más enérgico aplicando acciones privadas de disuasión y corrección. Ambos, Gould y Chomsky, han sido vigorosos proponentes del punto de vista de que los intelectuales son responsables de la aplicación y también de las erróneas aplicaciones de su propio trabajo, por lo que, presumiblemente, ellos están, al menos, disgustados por ser las fuentes citadas de todo este sinsentido, aunque ellos mismos no mantengan estos puntos de vista. (Quizá sea demasiado esperar su gratitud por hacerles este trabajo sucio).

3. *Buenos intentos*

> Al estudiar la evolución del entendimiento no podemos estimar hasta qué punto existen alternativas físicamente posibles para una gramática generativa transformadora en un organismo que reúna otras determinadas condiciones físicas características de los seres humanos. Es concebible que no exista ninguna —o muy pocas—, en cuyo caso hablar acerca de la evolución en la capacidad del lenguaje no viene a cuento.

Noam Chomsky, *Language and Mind*

Para progresar en la comprensión de todo esto, necesitamos probablemente comenzar con modelos simplificados (¿o hipersimplificados?) e ignorar el vehemente discurso de los críticos acerca de la mayor complejidad del mundo real. El mundo real es siempre más complejo, lo que tiene la ventaja de que no nos quedaremos sin trabajo.

John Ball, «*Mentes as Replications*»

El problema es cómo hacer que el péndulo detenga su balanceo tan destructivo. Una y otra vez observamos el mismo fallo de comunicación. El verdaderamente infortunado hiato de comunicación del que Simon y Kaplan hablan (en su cita al principio de la sección anterior) es el efecto amplificado de un relativamente sencillo *bit* de incomprensión inicial. Recordemos las diferencias entre reduccionistas y avariciosos reduccionistas (capítulo 3, sección 5): los reduccionistas creen que todas las cosas de la naturaleza pueden explicarse sin ganchos celestes; los reduccionistas avariciosos piensan que todo esto puede ser explicado sin grúas. Pero el optimismo saludable de un teórico es avaricia impropia en otro teórico. Un bando propone una grúa muy simplificada, de la cual se burlan en el otro lado del péndulo —«¡reduccionistas filisteos!»— declarando, honestamente, que la vida es bastante más complicada que todo eso. «¡Atajo de locos buscadores de ganchos celestes!», murmuran los simples reduccionistas, con una hiperreacción defensiva. Esto es lo que ellos murmurarían, respectivamente, si tuvieran la solución; pero si ambas partes tuvieran la solución podrían ser capaces de ver cuáles eran realmente los problemas y evitar por completo la falta de comunicación. Esta es mi esperanza.

¿Cuáles son los puntos de vista actuales de Chomsky? Si él no cree que el órgano del lenguaje se formó por selección natural, ¿cómo explica su complejidad? El filósofo de la biología Peter Godfrey-Smith [1993] recientemente dirigió su atención sobre todos los puntos de vista que mantienen, de un modo u otro, que «hay complejidad en el organismo en virtud de la complejidad en su medio ambiente». Dado que éste era uno de los temas preferidos de Herbert Spencer, Godfrey-Smith propone que califiquemos como «spenceriano» este punto de vista[141]. Spencer fue un darwiniano, aunque también se podría decir que Darwin era un spenceriano. En cualquier acontecimiento, la síntesis moderna es spenceriana en su núcleo y es el spencerianismo de esta ortodoxia el que es atacado más a menudo, de un modo u otro, por los rebeldes. Por ejemplo, Alfred Eigen y Jacques Monod

son spencerianos, por insistir en que la función molecular solamente puede ser especificada a través de la selección ambiental (capítulo 7, sección 2; capítulo 8, sección 3), mientras que la insistencia de Stuart Kauffman de que el orden emerge a pesar de la selección del medio ambiente expresa un reto antispenceriano (capítulo 8, sección 7). La negativa de Brian Goodwin a aceptar que la biología fuera considerada una ciencia *histórica* es otro ejemplo de antispencerianismo, dado que niega que las interrelaciones históricas con los medios ambientes más tempranos sean la fuente de las complejidades que se encuentran en los organismos. El breve coqueteo de Gould y Lewontin [1979] con un *Baupläne* «intrínseco», del que dependería todo el diseño del organismo, menos los pequeños ajustes, es otra cosa.

La sugerencia de Chomsky de que es la física y no la biología (o la ingeniería) la responsable de la estructura del órgano del lenguaje es, como hemos visto, pura doctrina antispenceriana. Esto explica su mala interpretación de mi amistosa sugerencia acerca de pasar la responsabilidad a la biología. Yo asumía, como un buen adaptacionista spenceriano, que «los genes son el canal a través del cual el medio ambiente habla», como escribió Godfrey-Smith, mientras que Chomsky prefería pensar que los genes, recibían su mensaje de alguna fuente de organización intrínseca, ahistórica, no relacionada con el medio ambiente, que podríamos llamar «física». Los spencerianos piensan que aunque hubiera tales «leyes de forma», solamente podrían imponerse sobre las cosas a través de uno u otro proceso de selección.

El pensamiento evolucionista es, precisamente, otro capítulo en la historia del pensamiento spenceriano frente al no spenceriano. El adaptacionismo es una doctrina spenceriana, como lo es el conductismo de Skinner, y así sucede, generalmente, en cualquier variante del empirismo. El *empirismo* es la teoría que sostiene que proporcionamos a nuestras mentes detalles que proceden todos del medio exterior, por la vía de la experiencia. El adaptacionismo es la teoría que sostiene que la selección que ejerce el medio externo conforma gradualmente los genotipos de los organismos, moldeándolos de tal modo que los fenotipos que ellos dirigen sean una adaptación casi óptima con respecto al mundo al que se han enfrentado. El conductismo es la tesis que sostiene que lo que Skinner [1953:129-41] denominó «el medio ambiente controlador» es el que «conforma» la conducta de los organismos. Ahora podemos comprobar que el famoso ataque de Chomsky contra Skinner estaba más dirigido contra el punto de vista spenceriano de Skinner de que el medio ambiente conformó el organismo, que contra las limitaciones del modelo de Skinner para explicar *cómo* se produjo esta conformación.

Skinner sostenía que *una simple repetición* del fundamental proceso darwiniano —condicionamiento del operante— podría ser responsable de toda inteligencia y de todo aprendizaje, no solamente en las palomas, sino en los seres humanos. Cuando los críticos insistían en que el pensamiento y el aprendizaje eran más, mucho más complicados que esta explicación, Skinner (y sus seguidores) olfateó ganchos celestes y definió a los críticos del conductismo como dualistas, mentalistas e ignorantes anticientíficos. Esta percepción de Skinner fue errónea; los críticos —al menos los mejores— sencillamente estaban insistiendo en que la mente estaba compuesta de muchas más grúas de lo que Skinner imaginaba.

Skinner fue un adaptacionista avaricioso que trataba de explicar todo el diseño (y el poder del diseño) de un solo golpe. La respuesta apropiada a Skinner debería haber sido: «¡Interesante ensayo, aunque todo esto resulta ser bastante más complicado de lo que usted piensa!». Y se debe decir sin sarcasmo, porque para Skinner *era* un ensayo interesante. Fue una gran idea, que inspiró (o provocó) medio siglo de obstinada experimentación y construcción de modelos, de los cuales se aprendió mucho. Es una ironía que fueran los repetidos fallos de *otra* rama del reduccionistno avaricioso, denominada por Haugeland [1985] «Good Old-Fashioned AI» o GOFAI (‘la vieja inteligencia artificial’), lo que realmente convenció a los psicólogos de que la mente era, efectivamente, un fenómeno de sobresaliente complejidad arquitectónica, bastante más complicado de como lo describe el conductismo. La intuición que sirvió de fundamento a la GOFAI fue el hallazgo de Turing de que un ordenador podía ser *indefinidamente complicado*, pero que todos los ordenadores pueden estar constituidos de partes simples. Mientras que las partes simples de Skinner eran emparejamientos aleatorios de parejas de estímulo-respuesta, sometidas por el medio ambiente, una y otra vez, a la presión selectiva de reforzamiento, las partes simples de Turing eran estructuras internas de datos, diferentes «estados en una máquina», que podían configurarse para responder indefinida y diferencialmente a muchos diferentes *inputs*, creando conductas *input-output* de cualquiera sofisticación imaginable. Cuáles de estos estados internos estaban especificados de modo innato y cuáles tendrían que ser revisados por la experiencia, era algo que quedaba abierto a la investigación. AI igual que Charles Babbage (véase la nota 13 del capítulo 8), Turing observó que la conducta de una entidad no necesitaba ser una simple función de su propia historia de estimulación, dado que podía haber acumulado grandes cantidades de diseño a lo largo de los eones, lo que le permitiría utilizar su complejidad interna como mediación de

sus respuestas. Esta introducción abstracta fue eventualmente complementada por los modeladores de la GOFAI con artificios de deslumbrante complejidad, que aún quedaban cómicamente lejos de producir un acto cognitivo al estilo humano.

Hoy, la ortodoxia dominante en la ciencia cognitiva sostiene que los viejos y simples modelos de percepción, aprendizaje, memoria, producción de lenguaje y comprensión del lenguaje son de un orden de magnitud demasiado simple, aunque aquellos modelos simples eran, a menudo, ensayos interesantes, sin los cuales estaríamos aún admirando lo simple que todo esto podría ser, después de todo. Esto da algún sentido a equivocarse a favor del reduccionismo avaricioso, al ensayar el modelo más simple, antes de sumergirse en las complejidades. La sencilla genética de Mendel fue un ensayo interesante y así fue como la bastante más compleja «genética del saco de alubias» llegó a manos de los genetistas de poblaciones, aun cuando, a menudo, han confiado en hipersimplificaciones retrospectivas tan extravagantes que Francis Crick tuvo la tentación de expulsarlos del campo de la ciencia. Los cristales de arcilla de Graham Cairns-Smith son un ensayo interesante, y el programa de Art Samuel para jugar a las damas en el ordenador era un ensayo interesante, demasiado simple, como ya hemos aprendido, pero en la vía correcta.

En los primeros días de los ordenadores, Warren McCulloch y W. H. Pitts [1943] propusieron una «neurona lógica» espléndidamente simple, a partir de la cual debían tejerse las «redes neurales» y, durante algún tiempo, pareció como si quizás hubiesen encontrado solución para la parte más difícil del problema del cerebro. Es cierto que antes de que hicieran su modesta proposición, los neurólogos estaban desesperadamente confusos acerca de cómo debería pensarse la actividad cerebral. Tan sólo hay que retroceder en el tiempo y leer, en los libros más especulativos de las décadas de los treinta y de los cuarenta, sus animosos y torpes forcejeos, para comprobar la enorme elevación que la neurociencia consiguió a partir de McCulloch y Pitts[142]. Ellos hicieron posible a pioneros como Donald Hebb [1949] y Frank Rosenblatt [1962], cuyos «Perceptrones» fueron, como Minsky y Papert [1969) pronto señalaron, un ensayo interesante, aunque demasiado simple. Ahora, varias décadas después, otra ola de ensayos interesantes, más complicados pero todavía útiles, haciendo ondear la bandera del conectivismo, están explorando áreas del espacio de diseño que no examinaron por sus ancestros intelectuales[143].

La mente humana es una admirable grúa, y para construirla y mantenerla funcionando y al día, ha sido necesaria una gran cantidad de trabajo de diseño. Este es el mensaje spenceriano de Darwin. De un modo u otro, la historia de los encuentros con el medio ambiente a lo largo de los eones (y durante los últimos diez minutos) ha conformado la mente que precisamente tengo en este momento. Parte de este trabajo debe haber sido realizado mediante selección natural y el resto por uno u otro proceso de generación y prueba de las características que ya hemos analizado en este capítulo. Nada de esto es mágico; nada de esto implica la existencia de un interno gancho celeste. Cualquier modelo que propongamos para estas grúas será seguramente demasiado simple desde uno u otro punto de vista, pero nos estamos acercando poco a poco, ensayando primero con las ideas más simples. Chomsky ha sido uno de los críticos más destacados de estos interesantes ensayos, descartando totalmente los procedentes de B. F. Skinner, desde expertos y rebeldes disidentes de la GOFAI como Herbert Simon y Roger Schank, hasta todos los conectivistas, y siempre ha estado seguro de que sus ideas eran, con mucho, demasiado simples, pero también ha puesto de manifiesto una hostilidad hacia la *táctica* de ensayar en modelos simples, la cual ha elevado innecesariamente la temperatura de los debates. Supongamos, con el fin de clarificar la discusión, que damos por probado que Chomsky pueda ver mejor que nadie que el entendimiento y el órgano del lenguaje, el cual desempeña un papel central en nuestra superioridad sobre las mentes de los animales, son estructuras de una sistematización y complejidad tal que, hasta la fecha, empequeñecen todos los modelos. Más razón entonces, piensa uno, para buscar una explicación evolucionista de estos brillantes dispositivos. Pero aunque Chomsky nos descubrió la estructura abstracta del lenguaje —la grúa que es más responsable de la elevación hasta su posición, más que todas las otras grúas de la cultura—, nos ha desanimado enérgicamente a considerarlo como una grúa. No es extraño que los que anhelan la existencia de ganchos celestes con frecuencia hayan aceptado a Chomsky como su autoridad.

Sin embargo, Chomsky no es el único candidato. Entre los buscadores de ganchos celestes, John Searle es el otro defensor favorito y, desde luego, no es un chomskiniano. En el capítulo 8 (sección 4) hemos visto que Searle defendió una versión de la tesis de John Locke sobre la primera mente, bajo el cartel de intencionalidad original. Según Searle, los autómatas (ordenadores o robots) no tienen intencionalidad real; como mucho tienen un mero *como* si fuera intencional. Además, la intencionalidad original o real no puede

componerse de, derivarse de —o, presumiblemente, descender de— un simple *como si* fuera intencional. Esto le crea un problema a Searle, ya que, mientras la inteligencia artificial dice que estamos *compuestos* de autómatas, el darwinismo dice que *descendemos* de autómatas. Es difícil negar lo primero si admitimos lo segundo; ¿cómo puede algo nacido de autómatas ser otra cosa sino un muy, muy, extravagante autómata? ¿Hemos conseguido, de algún modo, velocidad de escape para dejar atrás nuestra herencia de autómatas? ¿Hay algún umbral que marque el comienzo de la intencionalidad real? La original jerarquía de Chomsky, incluso para los autómatas más extravagantes, le permite trazar la línea mostrando que la *mínima complejidad* de un autómata capaz de generar las sentencias del lenguaje humano lo colocan en una clase especial, aunque siga siendo una especie de autómata, pero al menos de una clase avanzada. Esto no es bastante para Chomsky. Como ya hemos visto, se asentó bien en su posición y dijo, en efecto: «Sí, el lenguaje marca la diferencia, pero no tratemos de explicar cómo se diseñó el órgano del lenguaje. Es un monstruo prometedor, un regalo, nada que podamos explicar aún».

Una posición embarazosa la de Chomsky, si se quiere mantener: el cerebro es un autómata pero no podemos hacer ingeniería revertida. ¿Es éste, quizás, un error táctico? Según Searle, Chomsky dio muchos pasos antes de plantar los pies. Debería haber negado que el órgano del lenguaje tuviera una estructura que podía incluso ser descrita en los términos de un autómata. Pero mientras se dedicaba a explicar el procesamiento de la información y de las reglas y representaciones y transformaciones algorítmicas en el órgano del lenguaje, Chomsky había entregado un rehén a la ingeniería revertida. Quizá la herencia de Chomsky como ingeniero de la radio ha vuelto para rondarle como un fantasma:

> Específicamente, la evidencia de la gramática universal es bastante más simple por la siguiente hipótesis: hay, sin duda, un dispositivo para la adquisición del lenguaje (LAD: *language adquisition device*) innato en los cerebros humanos, y el LAD limita la forma de los lenguajes que los seres humanos pueden aprender. Hay, además, una explicación a nivel del *hardware* en lo que se refiere a la estructura del dispositivo, y hay un nivel funcional de explicación, que describe qué tipos de lenguajes puede adquirir el infante humano aplicando este mecanismo. No se añade ningún poder predictivo o explicatorio si se dice que hay, además, un nivel con profundas reglas inconscientes de la gramática universal y, naturalmente, yo he tratado de sugerir que esta postulación es incorrecta en todos los sentidos (Searle 1992:244-245).

Según Searle, la idea global del *procesamiento de la información* en el cerebro, descrita de modo abstracto en términos de algoritmos que exhiben

una neutralidad del sustrato, es incoherente. «Hay procesos neurofisiológicos brutos y ciegos, y hay consciencia, pero no hay nada más» (*ibidem*: 228).

Esto es ciertamente enfrentarse con valor a una dolorosa situación, y la misma con la que se enfrentó Chomsky, pero en una posición algo diferente. El dispositivo para la adquisición del lenguaje (LAD) ha evolucionado sin duda, y así lo hizo la consciencia (Searle 1992:887ss), pero Chomsky está seguro de que no hay esperanza para la ingeniería revertida. Sin embargo, Chomsky se equivoca al dar por probada incluso la coherencia de una descripción del proceso a nivel del autómata, porque esto abre la puerta a una «fuerte inteligencia artificial».

Si la posición de Chomsky en una resbaladiza ladera resulta difícil de mantener, la de Searle es aún más embarazosa[144]. Searle da por probado, como puede verse en el pasaje antes citado, que existe una historia «funcional» para ser contada acerca de cómo trabaja el cerebro en la adquisición del lenguaje. También puede contarse, concede Searle, una historia «funcional» acerca de cómo partes del cerebro llegan a juicios en profundidad o en distancia, sobre la visión. «*Pero a este nivel funcional no hay contenido mental de ninguna clase*» (Searle 1992:234; la cursiva es de Searle). Searle se plantea a sí mismo el siguiente contraargumento, completamente razonable, de los científicos cognitivos: «la distinción [entre discurso "funcional" y discurso de "contenido mental"] no marca diferencia en la ciencia cognitiva». Continuamos diciendo lo que siempre hemos dicho, y lo que siempre hemos hecho, sustituimos simplemente la palabra "funcional" por la palabra "mental" en estos casos». (Esto es de hecho lo que Chomsky ha hecho a menudo, replicando a esas críticas. Véase, por ejemplo, 1980). Para responder a este contraargumento, Searle (1992:238) está obligado a dar un paso atrás: no sólo no hay explicación a nivel del procesamiento de la información en el cerebro, dice; no hay realmente nivel «funcional» de explicación en biología:

> Para decirlo de manera directa, además de sus varias relaciones causales, el corazón no tiene funciones. Cuando hablamos de sus funciones, estamos hablando de aquellas de sus relaciones causales a las cuales concedemos alguna importancia *normativa*… En resumen, los hechos de intencionalidad real contienen elementos normativos, pero cuando nos referimos a las explicaciones funcionales, los únicos hechos son hechos físicos brutos y ciegos y las únicas normas están en nosotros, y existen solamente desde nuestro punto de vista.

Resulta entonces que hablar de función en biología, como mero discurso *como si* fuera intencional, no es realmente, después de todo, algo que deba ser tomado en serio. Según Searle, tan sólo los artefactos hechos por genuinos

artífices humanos conscientes tienen funciones *reales*. Las alas de los aviones son realmente para volar, pero las alas de las águilas no. Si un biólogo dice que hay adaptaciones para volar y otro dice que son simplemente unos espacios dispuestos para las decorativas plumas, la pregunta sobre qué biólogo está más cerca de la verdad no tiene sentido. Si, por otro lado, preguntamos a los ingenieros aeronáuticos si las alas de los aviones que ellos diseñan son para mantener el avión en vuelo o para disponer en ellas la insignia de la compañía aérea, puede que nos contesten con una barbaridad. De este modo, Searle termina negando la premisa de William Paley: según Searle, la naturaleza no consiste en una inimaginable variedad de dispositivos *funcionantes*, que exhiben diseño. ¡Sólo los artefactos humanos tienen este honor, y sólo a causa (como Locke nos «demostró») de que es a la mente a la que corresponde hacer algo con alguna función[145]!

Searle insiste en que las mentes humanas tienen intencionalidad «original», una propiedad inalcanzable, en principio, para cualquier proceso de I+D que pretenda construir cada vez mejores algoritmos. Esto es pura expresión de la creencia en los ganchos celestes: las mentes son fuentes de diseño originales e inexplicables, no resultados del diseño. Searle defiende su posición con más viveza que otros filósofos, aunque no está solo. La hostilidad hacia la inteligencia artificial y su demonio gemelo, el darwinismo, subyace justamente bajo la superficie de muchos de los trabajos más influyentes en la filosofía reciente del siglo XX, como veremos en el próximo capítulo.

CAPÍTULO 14
La evolución de los significados

I. *La búsqueda del significado real*

> —Cuando yo utilizo una palabra —dijo Humpty Dumpty, en un tono algo desdeñoso— significa precisamente lo que yo quiero que signifique, ni más ni menos.
>
> —La cuestión es —dijo Alicia— si tú *puedes* hacer que las palabras signifiquen muchas cosas diferentes.
>
> —La cuestión es —dijo Humpty Dumpty— quién es el que manda, eso es todo.
>
> Lewis Carrol, *A través del espejo*

No hay ningún problema filosófico que haya recibido más atención que el significado, en sus variadísimas manifestaciones. En el extremo principal del espectro los filósofos de todas las escuelas han tratado de resolver la cuestión fundamental del significado de la vida (y si esta cuestión tiene algún sentido o no lo tiene). En el extremo más modesto, los filósofos de la escuela analítica contemporánea —a veces denominada «filosofía lingüística» por los intrusos — han sometido los matices del significado de las palabras y la totalidad del lenguaje a un examen microscópico, en una variedad de proyectos completamente distintos. Retrocediendo a las décadas de los cincuenta y los sesenta, la escuela de la «filosofía del lenguaje ordinario» prodigó su atención a las sutiles diferencias entre palabras concretas: las diferencias entre hacer algo «deliberadamente» o «intencionalmente» o «con propósito», por citar un famoso ejemplo (Austin 1961). Todo esto abrió el camino a un conjunto de investigaciones más formales y sistemáticas. Qué diferentes proposiciones significa una sentencia como «Tom cree que Ortcutt es un espía».

¿Y qué teoría da razón de sus diferencias en presuposición, contexto e implicación? Este es el tipo de pregunta que ha buscado afanosamente la subescuela que ha sido a veces llamada «grupo de trabajo de la actitud preposicional», algunos de cuyos recientes esfuerzos ejemplares están en

Peacocke [1992] y Richard [1992]. Un diferente conjunto de investigaciones fue inaugurado por la teoría de Paul Grice [1957, 1969] acerca del «significado no natural». Esta teoría era un intento de especificar las condiciones bajo las cuales un *bit* de conducta no tiene precisamente significado natural (donde hay humo, hay fuego; —cuando alguien llora, allí hay tristeza), sino el tipo de significado que tiene el acto del habla con su elemento. ¿Cuál ha de ser el estado de la mente del que habla (o del que escucha) durante el acto de pronunciación para que signifique algo en general, o para que signifique una cosa en particular? O, en otras palabras, ¿cuál es la relación entre la psicología de un agente y el significado de las palabras de un agente? (Para las relaciones entre estos dos proyectos quizás sea mejor leer a Schiffer 1987).

Una hipótesis compartida por todos estos programas de investigación filosófica es la de que hay un tipo de significado —quizá dividido en mucho subtipos— que es lenguaje-dependiente. Antes de que hubiera palabras, no había significados de las *palabras*, incluso aunque hubiera otros tipos de significados. La ulterior hipótesis de trabajo, especialmente en los filósofos de lengua inglesa, ha sido que, hasta que consigamos comprender con claridad cómo las palabras pueden tener significado, es improbable que logremos progresos en las otras variedades de significados, especialmente en aquellos temas tan desconcertantes como el significado de la vida. Pero esta razonable hipótesis ha tenido un efecto colateral innecesario y debilitante: los filósofos, al concentrar su atención en primer lugar en el significado lingüístico, han distorsionado su visión de las mentes de las que estas palabras dependen, tratándolas como algo *sui generis*, más que como productos evolucionados del mundo natural. Este hecho se pone de manifiesto especialmente en la resistencia que los filósofos han mostrado ante las teorías evolucionistas del significado, teorías que proponen discernir que el significado de las palabras y de todos los estados mentales que, de algún modo, se encuentran tras aquellas, se fundamenta, en último término, en la fértil tierra de la función biológica.

Por un lado, pocos filósofos, si hay alguno, han pretendido negar un hecho obvio: los seres humanos son productos de la evolución, y su capacidad para hablar y, en consecuencia, para significar algo (en el sentido relevante del término), se debe a una serie de adaptaciones específicas no compartidas con otros productos de la evolución. Por otro lado, los filósofos se han mostrado poco dispuestos a aceptar la hipótesis de que el pensamiento evolucionista pueda arrojar luz a los problemas de cómo es que estas palabras, y sus fuentes y sus destinos en las mentes o cerebros humanos, poseen significado. Ha

habido importantes excepciones: Willard Van Orman Quine [1960] y Wilfrid Sellars [1963]; cada uno de ellos desarrolló teorías funcionalistas del significado, con razones basadas en la biología, de manera firme aunque escéptica. Sin embargo. Quine se comprometió con demasiada firmeza con el conductismo adoptado por su amigo B. F. Skinner, y ha sido acosado durante treinta años por intentar persuadir a los filósofos —con escaso éxito— de que sus tesis no caen bajo la universal denuncia de reduccionismo avaricioso lanzada por los cognitivistas en alza, bajo la dirección de Chomsky y Fodor[146], sobre Skinner y todos los conductistas. Sellars, el padre del «funcionalismo» en la filosofía de la mente, dijo todas las cosas correctas, aunque con un lenguaje difícil que fue ignorado en su mayor parte por los cognitivistas (véase Dennett 1987b:cap. 10, para una revisión histórica). Previamente, John Dewey dejó claro que el darwinismo debe asumirse como el fundamento de cualquier teoría naturalista del significado.

Ninguna explicación del universo en términos *simplemente* de redistribución de la materia en movimiento es completa, no importa hasta qué punto sea verdadera, si ignora el hecho cardinal de que el carácter de la materia en movimiento y su redistribución es acumulativo para conseguir unos fines; para crear el mundo de valores que conocemos. Si se niega esto se está negando la evolución; si se admite, estamos admitiendo un propósito en el único sentido objetivo —esto es, el único inteligible— de este término. No estoy diciendo que además del mecanismo haya otras causas o factores ideales que intervengan. Sólo insisto en que hay que contar la totalidad de la historia, que el carácter del mecanismo ha de ser anotado; es decir, que es de determinada manera para producir y sostener lo bueno en una multiplicidad de formas (Dewey 1910:34).

Nótese con cuánto cuidado Dewey sigue su camino entre Escila y Caribdis: no son necesarios ganchos celestes (causas o factores ideales), pero no debemos suponer que puede tener sentido una versión *no interpretada* de la evolución, una evolución sin funciones atribuidas, sin significados discernidos. Más recientemente, varios filósofos y yo mismo hemos articulado explicaciones específicamente evolucionistas del nacimiento y del mantenimiento del significado, tanto lingüístico como prelingüístico (Dennett 1969, 1978,1987b; Millikan 1984,1993; Israel 1987, Papineau 1987). La exposición de Ruth Millikan es, desde luego, la más cuidadosamente articulada, ya que no olvida los detalles de otros planteamientos filosóficos del significado arriba mencionados. Sus diferencias con mi posición han sido más amplias para ella que para mí, pero el hiato se está cerrando rápidamente (véase especialmente Millikan 1993:155) y espero que este libro lo acorte aún más, aunque éste no sea el lugar apropiado para exponer las diferencias que persisten, ya que son menores en el contexto de una mayor refriega, en una

batalla que todavía no hemos ganado: la batalla a favor de *cualquier* explicación evolucionista del significado.

Opuesto a nuestra posición se asienta un grupo de raras parejas, tan eminente como inverosímil: Jerry Fodor, Hilary Putnam, John Searle, Saul Kripke, Tyler Burge y Fred Dretske, contrarios, cada uno en lo suyo, tanto a una explicación evolucionista del significado como a la Inteligencia Artificial (IA). Estos seis filósofos han expresado sus reservas acerca de la IA, pero ha sido Fodor el más franco en su condena de los planteamientos evolucionistas del significado. Sus diatribas contra todo naturalista desde Dewey (véase especialmente Fodor 1990:cap. 2) son, a menudo, complemente cómicas. Como ejemplo, al ridiculizar mi punto de vista dice: «*Los ositos de peluche son artificiales, pero los osos reales son también artificiales*. Nosotros rellenamos a unos y la madre naturaleza rellena a los otros. La filosofía está *llena* de sorpresas» (1990:87)[147].

Las palabras, según Humpty Dumpty, consiguen de nosotros su significado, pero, ¿de dónde conseguimos nosotros el significado? Lo que ejercitan Fodor y estos otros filósofos es una preocupación por el significado *real* opuesto al significado *ersatz* ('sucedáneo', en alemán en el texto), intencionalidad *intrínseca* u *original*, opuesta a intencionalidad *derivada*. En su deseo de ridiculizar la idea de los organismos considerados como artefactos, Fodor nos hizo ver la perspectiva desde la cual puede revelarse, mediante un experimento mental, el fallo central de su planteamiento, un fallo compartido también por los otros cinco filósofos[148].

Imaginemos una máquina automática dispensadora de bebidas no alcohólicas, diseñada y construida en Estados Unidos y equipada con un dispositivo transductor estándar para aceptar o rechazar monedas de un cuarto de dólar. Denominemos a este dispositivo un *two-bitser* (*two-bits* = 25 centavos de dólar). Normalmente, cuando se inserta un cuarto de dólar en la máquina *two-bitser* se provoca en el programa un «estado» al que llamaremos *Q*, el cual «significa» (nótese el entrecomillado): «Yo percibo/acepto, en este momento, un genuino cuarto de dólar». Estas máquinas *two-bitser* son muy «listas» y «sofisticadas». (Más entrecomillado; el experimento mental comienza con la hipótesis de que este tipo de intencionalidad no es la real, y termina exponiendo las dificultades que tal hipótesis entraña). Sin embargo, estas máquinas a las que hemos denominado *two-bitser* difícilmente son infalibles: «cometen errores». Para decir lo mismo de manera no metafórica, a veces alcanzan el «estado» *Q* cuando una falsa moneda u otro objeto extraño se inserta en su ranura, y a veces rechazan monedas de 25 centavos

completamente legales, ya que en este caso fallan en el paso al «estado» *Q* cuando se supone que lo harían. Sin duda existen patrones detectables en los casos de «percepción equivocada». No hay duda de que, al menos en algunos casos, la «percepción equivocada» podría ser prevista por alguien con bastante conocimiento de las relevantes leyes de la física y de los parámetros del diseño de las máquinas *two-bitser*. En otras palabras, de las leyes de la física se deriva directamente que los objetos de la clase *K* colocarían al dispositivo de la máquina en el «estado» *Q*, del mismo modo que lo hacen los cuartos de dólar. Los objetos de la clase *K* serían buenas «monedas falsas» que volverían «loco» al transductor.

Si los objetos de la clase *K* llegasen a ser más frecuentes en el ambiente normal en el que se encuentran estas máquinas *two-bitser*, sería de esperar que sus propietarios y diseñadores desarrollaran transductores más avanzados y sensibles que discriminaran fiablemente entre los genuinos cuartos de dólar norteamericanos y los remedos del tipo K. Es natural que hicieran entonces su aparición las falsificaciones de los tramposos, lo que requeriría posteriores avances en la capacidad de detección de los transductores. En algún momento, esta escalada de ingeniería llegaría a provocar una disminución de la rentabilidad de las máquinas, ya que no existe un mecanismo infalible. Mientras tanto, los ingenieros y los usuarios son lo suficientemente listos como para seguir utilizando las rudimentarias máquinas *two-bitser*, tipo estándar, dado que es mucho más caro (relación coste-beneficio) intentar protegerse de abusos insignificantes.

La única cosa que hace que un dispositivo de un detector de un cuarto de dólar sea preferible a un detector de falsas monedas que simulan un cuarto de dólar o a un detector de un cuarto o de una falsa moneda, es el contexto de intenciones compartidas de los diseñadores, constructores, propietarios del artefacto; en resumen, de sus usuarios. Es solamente en el contexto de aquellos usuarios y de sus intenciones en el que podemos seleccionar algunas de las ocasiones en las que la máquina alcanza el «estado» *Q* como «verídicas» y en otras como «equivocadas». Sólo con relación a ese contexto de intenciones podemos justificadamente denominar al dispositivo, en primer lugar, como una máquina *two-bitser*.

Supongo que hasta este momento Fodor, Putnam, Searle, Kripke, Burge y Dretske coinciden conmigo: esto es justamente lo que sucede con tales artefactos; es, sencillamente, un caso de clara intencionalidad *derivada*. Tal artefacto no tiene, en modo alguno, intencionalidad *intrínseca*. De este modo, a nadie le crea un conflicto admitir que un dispositivo *two-bitser*, salido

directamente de una fábrica norteamericana, y con un letrero que diga «Modelo A *two-bitser*», pueda instalarse en una máquina dispensadora automática de bebidas no alcohólicas, donde proceda a mantener sus ganancias como un aceptador o rechazador de las monedas de un cuarto de balboa, moneda legal en Panamá, distinguiéndolas fácilmente de las monedas de cuarto de dólar de Estados Unidos, por el diseño y por las palabras estampadas, pero no por su peso, grosor, diámetro ni material. (No me lo estoy inventando. He sabido, por medio de una excelente autoridad en la materia —Albert Erler del Establecimiento de Monedas Raras «Flying Eagle»—, que las máquinas dispensadoras automáticas de tipo estándar no distinguen las monedas panameñas de un cuarto de balboa, acuñadas entre 1966 y 1984, de las monedas de cuarto de dólar. Una relativa sorpresa dado que fueron acuñadas a partir del stock de monedas de un cuarto de dólar que existía en las fábricas de monedas de Estados Unidos. Y, para satisfacer a los curiosos, aunque es estrictamente irrelevante en este ejemplo, el cambio de la moneda panameña, la última vez que cambié dólares, fue de un cuarto de dólar por balboa).

Esta máquina *two-bitser*, trasladada bruscamente a Panamá, estaría normalmente en un cierto «estado» físico —el «estado» con las características físicas que identificamos como «estado» *Q*—, sea un cuarto de dólar, un objeto metálico de la clase *K* o un cuarto de balboa panameño lo que se inserte en su ranura, aunque ahora un diferente grupo de casos se Kan de considerar como errores. En el nuevo ambiente, las monedas de cuarto de dólar cuentan como monedas falsas, como inductoras de error, falsa percepción equivocada, falsa representación, precisamente como hacen los objetos de la clase *K*. Después de todo, de regreso a Estados Unidos un cuarto de balboa panameño es un tipo de moneda falsa.

Una vez que nuestra máquina *two-bitser* reside en Panamá, ¿debemos decir que aún ocurre el «estado» que solemos llamar *Q*? El estado físico en el cual el dispositivo «acepta» monedas aún tiene lugar, pero ¿lo identificaríamos como «realización» de un nuevo «estado», al que llamaríamos *QB*? ¿En qué momento podríamos decir que el significado, o la función, de este «estado» físico de la máquina *two-bitser* ha cambiado? Bien, hay una gran libertad —por no decir hastío— acerca de lo que deberíamos decir, dado que, después de todo, una máquina *two-bitser* es justamente un artefacto; hablar acerca de sus percepciones y falsas percepciones, sus estados verídicos y no verídicos —en resumen, de su intencionalidad— son «metáforas» sin más. El estado interno de la máquina *two-bitser*; llámese

como se quiera, no significa *realmente* (originalmente, intrínsecamente) «aquí hay ahora un cuarto de dólar de Estados Unidos» o bien «aquí hay ahora un cuarto de balboa panameño». *Realmente* no significa nada. Esto es lo que Fodor, Putnam, Searle, Kripke, Burge y Dretske, entre otros, insistirían en decir.

La máquina *two-bitser* fue originalmente diseñada como un detector de monedas norteamericanas de un cuarto de dólar. Esta era su «función propia» (Millikan 1984) y, literalmente, su *raison d'être*. Nadie se hubiera molestado en construirla si no se les hubiera ocurrido este propósito. Este hecho histórico nos autoriza a llamar a esa cosa un *two-bitser*, una cosa cuya función es detectar cuartos de dólar, de modo que, *en relación con esta función*, podemos identificar tanto sus estados verídicos como sus errores.

Esto no impediría que la máquina *two-bitser* fuese arrebatada de su nicho original y presionada para realizar un nuevo servicio, con cualquier nuevo propósito que las leyes de la física permitan y las circunstancias favorezcan. Puede usarse como un detector *K*, o un detector de monedas falsas, un detector de monedas de un cuarto de balboa, el tope de una puerta o un arma mortífera. En su nuevo papel podría tener un breve período de confusión o indeterminación. ¿Cuánto tiempo ha de transcurrir en su nueva función para que ya no sea una máquina *two-bitser* sino una máquina detectora de monedas de un cuarto de balboa (podemos denominarlo un q-balboa)? Desde el mismo comienzo como máquina q-balboa, después de diez años de leal servicio como una máquina *two-bitser* ¿es el estado que se produce cuando se le presenta a la máquina un cuarto de balboa una verídica detección de un cuarto de balboa, o bien puede haber algo parecido a un error por la fuerza del hábito, un error por nostalgia, una errónea aceptación de un cuarto de balboa *como* un cuarto de dólar norteamericano?

Tal como la hemos descrito, la máquina *two-bitser* es demasiado simple como para que pueda tener una especie de memoria como la que nosotros tenemos de nuestras experiencias pasadas; sin embargo podemos intentar darle una. Supongamos que tiene un contador, el cual avanza cada vez que la máquina entra en un estado de aceptación, y éste señala, después de diez años de servicio, 1.435.792. Supongamos que el contador no se coloca a cero cuando es trasladada a Panamá, así que tras su inicial aceptación de un cuarto de dólar, se lee en el contador 1.435.793. ¿Inclina esto la balanza en favor de la afirmación de que la máquina todavía no se ha adaptado a la tarea de identificar cuartos de balboa? (Podemos ir añadiendo complicaciones y

variaciones si, según nuestra intuición, añaden diferencias. ¿Debemos hacerlo?).

Una cosa está clara: no hay absolutamente nada intrínseco acerca de la máquina *two-bitser* considerada minuciosamente en todos sus detalles (y en sus operaciones internas) que la distinga de una máquina detectadora de cuartos de balboa, construida por orden del Gobierno de Panamá. Lo que marca la diferencia, naturalmente, es sí ella fue *seleccionada* por su capacidad para detectar monedas de un cuarto de balboa (de acuerdo con Millikan 1984). Si fue así seleccionada (por sus nuevos propietarios, en el caso más simple), entonces, aunque ellos olvidaran resituar el contador, su movimiento inicial sería la aceptación verídica de un cuarto de balboa. «¡Esto funciona!», podrían exclamar encantados sus nuevos propietarios. Si, por otro lado, las máquinas *two-bitser* fueran enviadas a Panamá por error, o llegaran allí por pura coincidencia, entonces su primer movimiento no significaría nada, aunque pronto podría llegar a ser apreciado por aquellos de la vecindad debido a su poder de detectar cuartos de balboa entre las monedas falsas indígenas, en cuyo caso llegaría a funcionar como una máquina detectora de cuartos de balboa en el más completo sentido del término, a través de una vía menos oficial. Esto, dicho sea de paso, crea inmediatamente un problema al punto de vista de Searle, según el cual tan sólo los artefactos pueden tener funciones y éstas son las funciones con las que sus creadores les dotan mediante los muy especiales actos mentales de creación. Los diseñadores originales de la máquina *two-bitser* pueden no estar enterados de algún uso posterior para el cual fue oportunistamente adaptada, así que sus intenciones no cuentan para nada. Y los nuevos selectores pueden también fracasar en la *formulación* de algunas intenciones específicas; pueden caer en el hábito de confiar en la máquina *two-bitser* para alguna función a mano, sin darse cuenta del acto de exadaptación inconsciente que estaban ejecutando conjuntamente. Recordemos que Darwin, en *El origen de las especies*, llamó la atención sobre la selección inconsciente de características en los animales domésticos; la selección inconsciente de características en artefactos no es estirar las cosas: es más bien, cabe suponer, un acontecimiento frecuente.

Presumiblemente, Fodor y compañía no querrían estar en desacuerdo con este tratamiento de los artefactos, los cuales, afirman ellos, no tienen ni pizca de intencionalidad real; sin embargo pueden comenzar a preocuparse de que yo haya maniobrado hasta llevarles a un terreno resbaladizo. Pero permitidme ahora que considere un caso exactamente paralelo: «lo que los ojos de la rana dicen al cerebro de la rana». En su clásico artículo Lettvin, Maturana,

McCulloch y Pitt [1959] —otra obra maestra del Institute of Radio Engineers — demostraron que el sistema visual de la rana es sensible a pequeños puntos oscuros que se mueven ante la retina, mínimas sombras representadas casi siempre en circunstancias naturales por las moscas que vuelan en los alrededores. Este mecanismo «detector de moscas» está adecuadamente conectado al disparador en la lengua de la rana, lo que explica cómo las ranas se alimentan en un mundo cruel ayudando de esta manera a propagar su progenie. Entonces, ¿qué le *dice* el ojo de la rana al cerebro de la rana? ¿Que hay una mosca alrededor o que se trata de una falsa mosca (de un tipo o de otro) o una cosa de la clase *F* (de *fly*, 'mosca'), en suma, cualquier cosa capaz de disparar este mecanismo visual? Millikan, Israel y yo, como darwinianos teóricos de la significación, hemos discutido este caso, y Fodor coge al vuelo la oportunidad para mostrar dónde está el error, en su opinión, con cualquier planteamiento evolucionista de tales significados: son demasiado indeterminados. Fallan en la distinción, como debían hacerlo, entre los mensajes del ojo de la rana que dicen «aquí hay ahora una mosca» y «aquí hay ahora una mosca o un pequeño proyectil oscuro», y así sucesivamente. Pero esto es falso. Podemos utilizar el medio ambiente de selección de la rana (hasta la extensión que podamos determinar lo que ha sido) para distinguir entre los varios candidatos. Para hacer esto, usamos exactamente el mismo razonamiento que hemos utilizado para ordenar las interrogaciones —hasta la extensión que vale la pena ordenarlas— acerca del significado del estado en que se encuentra la máquina *two-bitser*. Y en tanto en cuanto no haya noticia acerca de lo que ha sido el medio ambiente de selección de la rana, la verdad es que no sabemos qué significa *realmente* el mensaje que transmite el ojo de la rana al cerebro de la rana. Esto puede demostrarse de modo concluyente enviando la rana a Panamá, o más precisamente, enviándola a la selección de un nuevo medio ambiente.

Supongamos que los científicos reúnen una pequeña población de ranas de alguna especie caza-moscas al borde de la extinción, y que las colocan bajo custodia protectora en un nuevo medio ambiente: un zoo especial para ranas donde no hay en absoluto moscas, aunque los cuidadores de este zoo periódicamente se disponen a lanzar a sus ranas alimentos preparados en pequeñas bolitas. Para delicia de los cuidadores, el procedimiento funciona; las ranas lanzan súbitamente sus lenguas hacia estas bolitas y después de algún tiempo existe multitud de ranas descendientes de éstas que nunca han visto en su vida una mosca, solamente han visto bolitas de alimentos. ¿Qué les dicen *sus* ojos a *sus* cerebros? Si insisten en decir que el significado no ha

cambiado, se encuentran en un aprieto porque esto es simplemente un claro ejemplo artificial de lo que continuamente sucede en la seleción natural: exadaptación. Como Darwin tuvo cuidado en recordarnos, la reutilización de maquinaria para nuevos propósitos es uno de los secretos del éxito de la madre naturaleza. Podemos reforzar el mensaje a cualquiera que desee persuasión adicional, suponiendo que de las ranas cautivas no todas lo hacen igualmente bien, debido a variaciones en la capacidad de sus ojos en detectar las bolitas, algunas comen con menos motivación que otras y, como resultado, su progenie es menor. Enseguida se habrá producido una innegable selección para la detección de las bolitas, aunque sería un error preguntar exactamente en qué número ha ocurrido esto, para «contarlo».

A menos que hubiera variación «sin significado» o «indeterminada» en las condiciones de disparo de los ojos de varias ranas, no habría materia prima (variaciones ciegas) para que la selección actuara con un *nuevo* propósito. La indeterminación que Fodor (y otros) considera como un fallo en la explicación darwiniana de la evolución del significado es realmente una precondición para tal evolución. La idea de que debe ser *algo determinado* lo que el ojo de la rana realmente significa —alguna posiblemente irreconocible proposición en el lenguaje de la rana que expresase exactamente lo que el ojo de la rana está diciéndole al cerebro de la rana— es justamente esencialismo aplicado a la significación (o función). La significación, como la función, de la cual tan directamente depende, no es algo determinado en el momento de su nacimiento. No surge mediante un salto o por creación especial, sino por una modificación (característicamente gradual) de las circunstancias.

Ahora estamos listos para la única cuestión que realmente interesa a estos filósofos: ¿qué sucede cuando trasladamos a una *persona* de un medio ambiente a otro? Existe el conocido experimento mental de la Tierra Gemela de Hilary Putnam [1975]. Me resisto a entrar en detalles, pero he aprendido que tan sólo una lectura laboriosa y completa, más el bloqueo de todas las salidas, tendrá un efecto persuasivo para aquellos cuyas fidelidades se mantienen con intencionalidad original. Así que, con excusas, en esas estamos. Armados con los fundamentos de nuestro análisis sobre la máquina expendedora de bebidas (*two-bitser*) y sobre la rana, podemos ver exactamente cuánto depende el experimento mental de la Tierra Gemela de su innegable fuerza retórica. La Tierra Gemela, vamos a suponerlo, es un planeta casi exactamente igual a la Tierra, excepto que allí no hay caballos. Hay animales que parecen caballos a los cuales los habitantes del Boston Gemelo y del Londres Gemelo llaman *horses*; también llamados *chevaux* por los

habitantes del París Gemelo, y así sucesivamente, del mismo modo que la Tierra Gemela es similar a la Tierra. Pero estos animales de la Tierra Gemela no son caballos, son otra cosa. Les podemos llamar *paracaballos*, y se puede suponer que son como pseudomamíferos, reptiles peludos o cualquier otra cosa; esto es filosofía y podemos añadir cualquier detalle que necesitemos para que el experimento mental «funcione». Ahora llega el momento dramático. Una noche, mientras un habitante de la Tierra (un terrícola) duerme, es trasladado súbita y mágicamente a la Tierra Gemela. (Es importante que esté dormido durante este cambio momentáneo, para que se mantenga a ciegas con respecto a lo que le ha sucedido, mantenido «en el mismo estado» en el que se encontraba en la Tierra). Cuando se despierta mira por la ventana y ve a un paracaballo galopando. «¡Mirad, un caballo!», dice (a gritos o susurrando, da igual). Para simplificar la historia, supongamos que Twinning, un habitante de la Tierra Gemela, emite el mismísimo sonido en el mismísimo momento en el que ve también al paracaballo galopando. Este es el punto en el que Putnam y otros insisten: Twinning dice y cree algo *verdadero* —es decir, que un paracaballo está precisamente galopando— mientras que el «terrícola» dice y cree algo *falso*; es decir, que lo que está corriendo es un caballo. ¿Cuánto tiempo, no obstante, tendrá que vivir en la Tierra Gemela el terrícola llamando caballos a los paracaballos (como todos los nativos), hasta que el estado de su *mente* (o de lo que sus ojos le dicen a su cerebro) sea una verdad acerca de los paracaballos y no una falsedad sobre los caballos? (¿Cuando «en lo que concierne» a la intencionalidad dé un salto a la nueva posición?, salto que exigen estos teóricos). ¿Hará el terrícola en algún momento esta transición? ¿Lo hará de alguna manera sin darse cuenta? Después de todo, la máquina *two-bitser* nunca tuvo noticia del cambio de significado de su estado interno.

Supongo que el terrícola que ha ido a parar a la Tierra Gemela puede optar por pensar que él es radicalmente distinto de la rana y de la máquina *two-bitser*. Este terrícola, al parecen posee intencionalidad intrínseca u original, y esta maravillosa propiedad tiene un cierto grado de inercia: su cerebro no puede devolver una moneda y súbitamente significa algo enteramente nuevo con relación a su viejo estado. En contraste, las ranas no tienen bastante memoria y la máquina *two-bitser* no tiene ninguna. Lo que *ese terrícola trasladado a la Tierra Gemela* entiende por la palabra «caballo» (su particular concepto mental de un caballo) es algo así como *una de esas bestias equinas que a nosotros, terrícolas, nos gusta montar*, un epíteto afincado en su mente a través de todas las exhibiciones de caballos y las

películas del Oeste. Estamos de acuerdo en que esta matriz de la memoria *fija* la clase de cosa a la cual su concepto de caballo se aplica. *Ex hypothesis* los paracaballos de la Tierra Gemela no son bestias de esta clase: en modo alguno son de la misma especie, aunque son, por conveniencia, indistinguibles de los caballos para el terrícola. Así, de acuerdo con esta línea de pensamiento, el terrícola cometerá inconscientemente errores cada vez que clasifique equivocadamente un paracaballo al verlo o al evocarlo. (¿No era un hermoso caballo lo que vi galopar ayer a través de mi ventana?).

Pero hay otra forma de ver el mismo ejemplo. Nada nos obliga a suponer que el concepto de caballo del terrícola no era más reposado en primer lugar que su concepto de una mesa. (Intente contar la historia de la Tierra Gemela con la sugerencia de que las mesas no eran realmente mesas, pero que aparecían como mesas y eran usadas como mesas. ¿A que no funciona?). Los caballos y los paracaballos pueden no ser de la misma especie biológica, pero qué será si uno, al igual que la mayoría de los habitantes de la Tierra, no tiene un concepto claro de las especies y las clasifica por su apariencia: *cosa viva parecida a Man-o-War* (famoso caballo de carreras de los años veinte). Los caballos y los paracaballos caen ambos dentro de esta clase, así que cuando el terrícola en la Tierra Gemela denomina a un animal como un caballo, está en lo cierto, después de todo. Dado lo que el terrícola entiende por «caballo», los paracaballos son caballos, una clase de caballo no terrícola, pero igualmente un caballo. Las mesas no terrícolas son también mesas. Está claro que el terrícola puede tener este concepto laxo de los caballos, mientras que para otro habitante de la Tierra el concepto puede ser más estricto, al entender que los paracaballos no son caballos, no son seres de la misma especie de los de la Tierra. Ambos casos son posibles. ¿Debe *determinarse* si el concepto de caballo del terrícola (antes de su traslado a la Tierra Gemela) significa la especie o una clase más amplia? Podría ser si el terrícola estuviera, por ejemplo, en posesión de suficientes conocimientos biológicos, pero supongamos que no los tiene. Entonces, el concepto del terrícola —lo que «caballo» *realmente* significa para él— sufrirá la misma indeterminación que el concepto de *mosca* para la rana (¿o acaso el concepto fue siempre: *pequeño objeto de alimento aerotransportado*?).

La siguiente historia puede ayudar a conseguir un ejemplo más realista, al tratarse de algo que podría suceder aquí mismo en la Tierra. Se cuenta que los siameses poseían una palabra para el «gato», pero nunca habían visto o imaginado otros gatos que no fueran siameses. Supongamos que la palabra fuese «kato»; no importan los detalles reales o si esta historia. Podría ser.

Cuando descubrieron que existían otras variedades, tuvieron un problema: ¿su palabra significaba «gato» o «gato siamés»?, ¿habían descubierto que existían otros tipos de «katos» de aspecto bastante diferente, o que estos «katos» y aquellas otras criaturas pertenecían a un supergrupo?, ¿era el término tradicional el nombre de una especie o de una variedad? Si no poseían la teoría biológica que establece esta distinción, ¿cómo era posible un hecho de esta naturaleza? (Bueno, los siameses podrían descubrir que las peculiaridades de la apariencia eran realmente muy importantes para ellos. ¡Esto justamente no se parece a un «kato», así que no lo es! Y de modo similar podemos descubrir que nos resistimos a la sugerencia de que los ponies de la raza Shetland sean caballos).

Cuando un siamés vio un gato (no siamés) caminando pausadamente y pensó «¡Mira un "kato"!», ¿era un error o la simple verdad? Quizás el siamés no tendría una opinión acerca de cómo responder a esta cuestión, pero, no obstante, ¿podría ser un hecho determinado acerca de si esto era un error, algo que nosotros nunca seremos capaces de describir, aunque, con todo, sea un hecho?

Después de todo, podía habernos sucedido lo mismo: imaginemos que un biólogo nos dice un día que los coyotes son, de hecho, perros, es decir, miembros de la misma especie. Nos podríamos preguntar si ambos tenemos el mismo concepto de perro. ¿Hasta qué punto es fuerte nuestra fidelidad al punto de vista de que el «perro» es el término que designa una especie y no el nombre para una amplia subespecie de perros *domésticos*? ¿Nos dice nuestra intuición, alto y claro, que ya hemos desechado «por definición» la hipótesis de que los coyotes sean perros, o bien con su silencio admite tácitamente que nuestro concepto ha tenido en todo momento la imparcialidad suficiente para admitir este supuesto descubrimiento? ¿O admitiremos, ahora que se ha planteado la cuestión, que tendríamos, por una vía o por otra, que adoptar una posición, algo que simplemente no había sido fijado antes debido a que nunca se había planteado?

Esta amenaza de indeterminación socava la argumentación del experimento mental de Putnam. Para preservar su tesis, Putnam trata de taponar la brecha declarando que nuestros conceptos —tengamos o no conocimiento de ellos— se refieren a los *tipos naturales*. Pero ¿qué tipos son naturales? Las variedades son tan naturales como las especies, y éstas tan naturales como los géneros y las clasificaciones superiores. El esencialismo respecto al significado del estado mental de la rana y el «estado interno» *Q* (o *QB*) se ha evaporado, para nosotros debe evaporarse con la misma seguridad.

La rana, libre en su ambiente natural, *podría* localizar con enorme rapidez y destreza las bolitas —tiene algunas en su camino— ya que no está ciertamente equipada con algo que discrimine en *contra* de las bolitas. En un sentido, *mosca* o *bolita* es un tipo natural para las ranas; éstas, *naturalmente*, fallan en la discriminación entre las dos. En otro sentido, *mosca o bolita* no es un tipo natural para las ranas; su medio ambiente natural nunca ha hecho antes esta clasificación relevante. Exactamente lo mismo es verdad para el terrícola catapultado súbitamente a la Tierra Gemela. Si los paracaballos hubiesen sido traídos secretamente a la Tierra desde la Tierra Gemela, el terrícola les habría llamado inmediatamente «caballos». Se habría equivocado si, de algún modo, se hubiera establecido que lo que su término significa era la especie, no lo parecido a la especie, aunque si no, no tendría en absoluto fundamentos para calificar a su clasificación como un error, dado que esta distinción nunca se habría establecido. El terrícola en la Tierra Gemela, al igual que la rana y la máquina *two-bitser*, tiene estados internos que derivan su significado de sus roles funcionales, y donde la función falla en la generación de una respuesta, no hay nada más que investigar.

La historia de la Tierra Gemela, si la leemos a través de lentes darwinianas, prueba que los significados humanos son exactamente tan derivados como los significados de la máquina *two-bitser* y los de las ranas. Esto no es lo que se intentaba demostrar, aunque cualquier intento de bloquear esta interpretación se ve forzado a postular misteriosas e inmotivadas doctrinas esencialistas y a insistir, directamente, en que hay hechos acerca del significado que son absolutamente inertes e indescubribles, pero, al fin y al cabo, hechos. Dado que algunos filósofos están dispuestos a tragarse estas amargas píldoras, es necesario que añada otras persuasiones.

La idea de que nuestros significados son exactamente tan dependientes de la función como los significados de los estados de los artefactos y, en consecuencia, igualmente derivados y potencialmente indeterminados, ha sido considerada intolerable por algunos filósofos, ya que falla en dar al significado un digno *rol causal*. Esta es una idea que ya hemos visto en una más temprana encarnación, como la preocupación porque las mentes sean meros *efectos*, no *causas* originantes. Si los significados están determinados por las fuerzas selectivas que apoyan ciertos roles funcionales, entonces todos los significados pueden parecer, en un sentido, sólo atribuidos *retrospectivamente*: lo que algo significa no es una propiedad *intrínseca*, capaz de establecer una diferencia en el mundo en el momento de su nacimiento, sino, cuanto más, una coronación retrospectiva asegurada

solamente por un análisis de los subsiguientes efectos generados. Esto no es completamente correcto: un análisis de la máquina *two-bitser* realizado por un ingeniero, nuevamente trasladada a Panamá, permitiría decir para qué roles merecería la pena utilizarse el dispositivo así configurado, aun cuando no hubiese sido escogida para ningún rol. Podemos alcanzar este veredicto: su estado de aceptación puede significar «un cuarto de balboa aquí y ahora» si la colocamos en el ambiente adecuado. Pero, naturalmente, esto puede significar también otras muchas cosas, si se coloca la máquina en otros medios ambientes, de modo que ninguna de estas máquinas adquiere un significado hasta que un concreto rol funcional le ha sido establecido, rol para cuyo establecimiento no hay umbral ni límite de tiempo.

Todo esto no basta para algunos filósofos que piensan que el significado, así construido, no tiene peso. La expresión más clara de esta idea es la insistencia de Fred Dretske [1986] en que el mismo significado debe tener una función causal en nuestras vidas mentales, mientras que en la carrera de un artefacto el significado nunca tiene esta función causal. Así planteada la cuestión, el intento de distinguir el significado real del significado artificial descubre inadvertidamente una verdadera obsesión por los ganchos celestes, una obsesión por algo basado en «elevados principios», que podrían bloquear la gradual emergencia de significado a partir de alguna cascada de meras causas mecánicas, sin propósito, pero esto es (como el lector sospechará) un modo opcional y tendencioso de presentar la cuestión. Como es habitual, los problemas son más complejos que como los he presentado[149], aunque podemos poner al descubierto los puntos clave con la ayuda de una pequeña fábula que recientemente he diseñado precisamente para estas pataletas de los filósofos. Funciona. Primero la fábula, y después las pataletas.

2. *Dos cajas negras*[150]

Había una vez dos grandes cajas negras, A y B, conectadas mediante un largo alambre de cobre, bien aislado. En la caja A había dos botones, marcados como α y β, y en la caja B había tres luces: roja, verde y ámbar. Los científicos que estudiaron la conducta de las cajas observaron que cada vez que se apretaba el botón α en la caja A, la luz roja parpadeaba brevemente en la caja B, y cada vez que se apretaba el botón β en la caja A, la luz verde parpadeaba en la caja B. La luz ámbar nunca parecía parpadear. Los científicos llevaron a cabo miles de millones de pruebas, bajo una amplia

variedad de condiciones, y no hallaron excepciones. Parecía existir una regularidad causal, que resumieron como sigue:

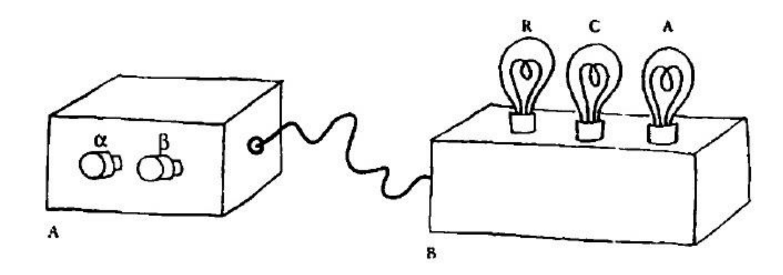


Figura 14.1

Todas las α causan algo en las luces rojas
Todas las β causan algo en las luces verdes

Los científicos determinaron que la causalidad pasaba, de algún modo, a través del alambre de cobre, dado que al cortarlo se eliminaban todos los efectos en la caja B y aislando dos cajas una de otra, sin cortar el alambre, nunca se interrumpía la regularidad. Así que era natural que los científicos mostraran curiosidad por conocer exactamente cómo funcionaba a través del alambre la regularidad causal que habían descubierto. Quizá, pensaron, al presionar el botón α se causaba un pulso de bajo voltaje que era emitido por el alambre, disparando la luz roja, y al presionar el botón β se provocaba un pulso de alto voltaje, que disparaba la luz verde. O, quizás al presionar α se causaba una simple pulsación que disparaba la luz roja y la presión β causaba un doble pulso. Era obvio que había algo que siempre sucedía en el alambre cuando se presionaba el botón α, y algo diferente que siempre sucedía cuando se presionaba el botón β. Descubrir justamente lo que esto era explicaría la regularidad causal que habían descubierto.

Un interruptor colocado en el alambre de cobre pronto reveló que las cosas eran más complicadas. Cualquiera que fuera el botón presionado en la caja A, una larga corriente de pulsos e hiatos (*on* y *off*, o *bits*) era enviada rápidamente por el alambre a la caja B —diez mil *bits*, para ser exactos—. ¡Pero cada vez el patrón de la corriente era distinto!

Parecía evidente que debía de ser una característica o propiedad de las tiras de *bits* lo que iluminaba la luz roja en un caso y la luz verde en el otro. ¿Qué sería? Decidieron abrir la caja B y ver lo que sucedía cuando llegaban las tiras de *bits*. Dentro de la caja B encontraron un superordenador, exactamente un superordenador digital ordinario, de serie, con una gran

memoria, que contenía un inmenso programa y una inmensa base de datos, escrita, naturalmente, en más tiras de *bits*. Y cuando los científicos siguieron la pista a los efectos de las tiras de *bits* que llegaban al programa del ordenador, no encontraron nada extraordinario: la tira de *inputs* parecía seguir siempre su camino hacia la unidad central de procesamiento, de manera normal, y allí provocaba unos cuantos miles de millones de operaciones, realizadas en unos cuantos segundos, terminando siempre, con una o dos señales de *output*, un I (que encendía la luz roja) o un O (que encendía la luz verde). En cada caso los científicos descubrieron que podían explicar al nivel microscópico, sin dificultad ni controversia, cada paso de la causalidad. No sospecharon la intervención de causas ocultas y, por ejemplo, cuando se dispusieron las cosas para recibir (*input*) la misma secuencia de diez mil *bits* una y otra vez, el programa en la caja B siempre generaba el mismo *output*, rojo o verde.

Pero esto era algo enigmático, ya que mientras que siempre daba como resultado el mismo *output*, no siempre generaba el mismo *output* pasando por los mismos pasos intermedios. En efecto, casi siempre pasaba a través de diferentes estados antes de generar el mismo *output*. Esto, en sí, no era un misterio, dado que el programa guardaba una copia de cada *input* recibido y así, cuando el mismo *input* llegaba una segunda o tercera vez, o mil veces, el estado de la memoria del ordenador era ligeramente diferente cada vez. Pero el *output* era siempre el mismo; si la luz era roja la primera vez, el *input* era una particular tira, y consiguientemente siempre sería roja la luz para la misma tira de *bits*, y la misma regularidad se verificaba en las tiras verdes (como los científicos comenzaron a llamar a las que encendían la luz de este color). Los científicos tuvieron la tentación de establecer la hipótesis de que todas las tiras eran, o bien rojas (causan una luz roja parpadeante), o bien verdes (causan una luz verde parpadeante). Pero, naturalmente, no habían probado todas las tiras de *bits* posibles; solamente las emitidas por la caja A.

Así que decidieron probar la hipótesis desconectando temporalmente A de B e insertando variaciones en las tiras de *outputs* de A a B. Pero, para su confusión y decepción, descubrieron que ¡casi siempre que manipulaban una tira de *bits* procedente de A era la luz *ámbar* la que parpadeaba! Era como si la caja B hubiera detectado la manipulación de la tira. No había duda, sin embargo, de que la caja B aceptaba sin dificultad versiones de tiras rojas fabricadas por humanos iluminando la luz roja y versiones de tiras verdes fabricadas por humanos iluminando la luz verde. Solamente cuando un *bit* —o más de un *bit*— se cambiaba en una tira roja o verde, generalmente —no

siempre— aparecía la luz ámbar. «¡Me la he cargado!», dijo alguien impulsivamente, tras observar una tira roja que había manipulado, convertida en una tira ámbar; y esto condujo a una racha de especulaciones acerca de que la tiras roja y verde estaban vivas, en algún sentido —quizá macho y hembra —, mientras las tiras ámbar eran tiras *muertas*, Pero esta atrayente hipótesis no conducía a ninguna parte, aunque una posterior experimentación con unos cuantos miles de billones de variaciones aleatorias sobre tiras de *bits* de diez mil *bits* de longitud, sugerían fuertemente a los científicos que había realmente tres variedades de tiras: tiras rojas, tiras verdes y tiras ámbar; las tiras ámbar sobrepasaban con mucho en número, en muchos órdenes de magnitud, a las rojas y verdes. Casi todas las tiras eran de color ámbar. Esto hacía que la regularidad que los científicos habían descubierto fuera aún más excitante y enigmática.

¿Qué había en las tiras rojas que encendía la luz roja y en las tiras verdes que encendía la luz verde? Naturalmente, en cada caso específico no había misterio. Los científicos podían seguir las huellas de la causalidad de cada tira específica a través del superordenador situado en la caja B y observar, con gratificante determinismo, que producirían luz roja, verde o ámbar, según el caso. Sin embargo, lo que no podían encontrar era un modo de predecir cuál de los tres efectos produciría una nueva tira de *bits*, examinándola con precisión (sin «estimularla manualmente» en la caja B). Los científicos sabían, a partir de datos empíricos, que eran muy altas las probabilidades de que una nueva tira fuera ámbar, a menos que fuera una tira que se sabía que había sido emitida por la caja A, en cuyo caso las probabilidades eran mejores de mil millones a uno de que sería o roja o verde, pero nadie podía decir de qué color sería, sin hacerla pasar por la caja B, para ver lo que establecía el programa.

Dado que, a pesar de la brillante y costosa investigación, los científicos se veían todavía absolutamente incapaces de predecir si una tira sería roja, verde o ámbar, algunos teóricos tuvieron la tentación de denominar *emergentes* a estas propiedades. Lo que con este término querían decir era que las propiedades eran —pensaban ellos— *impredecibles por principio* a partir de un mero análisis de las micropropiedades de las mismas tiras. Pero esto no parecía probable en absoluto, dado que cada caso concreto era tan predecible como sería cualquier *input* determinista en un programa determinista. En cualquier caso, fueran o no las propiedades de rojo, verde y ámbar impredecibles por principio o meramente en la práctica, ciertamente eran propiedades sorprendentes y misteriosas.

Quizá la solución de este misterio se encontraba en la caja A. Los científicos abrieron esta caja y encontraron otro superordenador, de diferente hechura y modelo, que hacía funcionar un programa gigantesco diferente, pero también un ordenador digital común. Pronto determinaron que cada vez que se presionaba el botón a éste apagaba el programa de un modo, con el envío de un código (IIIIIII) a la Unidad Central de Programación (UCP), y cada vez que se presionaba el botón p éste enviaba un código diferente (OOOOOOOO) a la UCP, poniendo en marcha un conjunto diferente de miles de millones de operaciones. El resultado era un «reloj» interno marcando millones de veces por segundo, y cada vez que se presionaba cualquier botón, lo primero que hacía el ordenador era tomar el «tiempo» del reloj (por ejemplo, IOIIOIOIOIOIOIOIII) y desintegrarlo en tiras que entonces usaría para determinar qué subrutinas serían reclamadas y en qué orden, y a qué parte de la memoria se accedería en primer lugar durante la preparación de una tira de *bits* que iba a ser enviada a través del alambre de cobre que unía las dos cajas.

Los científicos fueron capaces de deducir que era la consulta del reloj (prácticamente aleatoria) la que virtualmente garantizaba que la misma tira de *bits* no se enviara nunca dos veces. Pero, a pesar de esta aleatoriedad o pseudoaleatoriedad, seguía siendo verdad que cada vez que se presionaba el botón a la tira de *bits* que el ordenador confeccionaba resultaba ser la roja, y cada vez que se presionaba el botón β, la tira de *bits* que el ordenador enviaba era la verde. En realidad, los científicos encontraron algunos casos anómalos: en aproximadamente uno de cada mil millones de pruebas, la presión sobre el botón a causaba la emisión de una tira verde y al presionar el botón β se causaba la emisión de una tira roja. Este pequeño defecto en la perfección no hizo más que estimular el apetito de los científicos por encontrar una explicación de la regularidad.

Y entonces, un buen día, llegaron los dos *hackers* de inteligencia artificial (AI) que habían construido las cajas y lo explicaron todo. (No lea lo siguiente si desea aclarar por sí mismo el misterio). Al, que había construido la caja A, había estado trabajando durante años en un «sistema experto», una base de datos que contiene «proposiciones verdaderas» acerca de todo lo que existe bajo el sol, y una máquina de inferencias para deducir posteriores implicaciones de los axiomas que componían la base de datos. Eran estadísticas de la liga mayor de béisbol, informes meteorológicos, taxonomías biológicas, historias de las naciones del mundo, y una gran cantidad de hechos triviales de la base de datos. Bo, el sueco que había construido la caja

B, había estado trabajando simultáneamente en una base de datos rival sobre el conocimiento del mundo para su propio sistema experto. Habían llenado hasta los topes sus respectivas bases de datos con tantas «verdades» como los años de trabajo le habían permitido[151].

Pero con el paso de los años llegaron a aburrirse de los sistemas expertos y decidieron que las perspectivas prácticas de esta tecnología habían sido enormemente sobrevaloradas. Los sistemas no eran muy buenos para resolver problemas interesantes, o «pensar» o «encontrar soluciones creativas a problemas». Todos estos sistemas eran buenos cuando, gracias a sus máquinas de inferencia, generaban enormes cantidades de verdaderas frases (en sus respectivas lenguas) y sometían a prueba *inputs* de frases (en sus respectivas lenguas) sobre la verdad y la falsedad de sus «conocimientos». Así, Al y Bo habían calculado conjuntamente de qué modo podían utilizarse los frutos de sus esfuerzos baldíos. Decidieron construir un juguete filosófico. Eligieron una *lingua franca* para la traducción entre sus dos sistemas representativos (fue realmente el inglés, emitido en el código estándar ASCII, el código del correo electrónico) y conectaron las máquinas con un alambre. Cada vez que se presionaba el botón a de A, éste instruía a A para escoger aleatoriamente o pseudoaleatoriamente una de sus «creencias» (cualquier axioma almacenado o una implicación generada a partir de sus axiomas), ésta se traducía al inglés (los caracteres en inglés, en un ordenador, estarían ya en el sistema ASCII), se añadían bastantes *bits* aleatorizados durante el período necesario para elevar la suma total a diez mil *bits*, y se enviaba la tira resultante a la caja B, la cual traduciría este *input* a su propio idioma (que era el sueco) y se probaba contra sus propias «creencias» en su base de datos. Dado que ambas bases de datos estaban compuestas de verdades, y aproximadamente de las mismas verdades, gracias a las máquinas de inferencias, cada vez que A envía a B algo que A «cree», lo «cree» también B y lo señala con una parpadeante luz roja. Cada vez que A envía a B lo que A estima como una falsedad, B anuncia, parpadeando una luz verde, que ha juzgado que esto es una falsedad.

Y cada vez que alguien manipulaba la transmisión, esta acción casi siempre resultaba en una tira que no era una frase del inglés bien construida (B tiene una tolerancia que es absolutamente cero para errores «tipográficos»), B respondía con el parpadeo de una luz ámbar. Cada vez que alguien escogía aleatoriamente una tira de *bits* las probabilidades de que esta tira no presentara una verdad o una falsedad bien construida en el inglés

ASCII eran inmensas; en consecuencia, se produciría la preponderancia de las tiras de color ámbar.

De este modo, dicen Al y Bo, la propiedad emergente *rojo* era realmente la propiedad de ser una verdadera frase del inglés, y *verde* era la propiedad de ser una falsedad en inglés. De repente, la búsqueda que había ocupado a los científicos durante años se convertía en un juego de niños. Cualquiera podría componer tiras rojas *ad nauseam*, por ejemplo, solo con escribir el código ASCII para «Las casas son más grandes que los cacahuetes» o «Las ballenas no vuelan» o «Tres veces cuatro es dos veces menos que dos veces siete». Si desea una cinta verde, intente escribir «Nueve es menos que ocho» o «Nueva York es la capital de España».

Los filósofos pronto descubrieron trucos, tales como encontrar tiras que fueran rojas las primeras cien veces que fueran enviadas a B, pero verdes de ahí en adelante (por ejemplo, la versión en código ASCII de «Esta sentencia le ha sido enviada para evaluación sólo hasta cien veces»).

No obstante, dicen algunos filósofos, las propiedades de las tiras *roja* y *verde* no son realmente *verdaderas en inglés y falsas en inglés*. Después de todo, son verdades en inglés cuya expresión en código ASCII necesita de millones de *bits* y, además, a pesar de sus esfuerzos, Al y Bo no siempre insertaban *hechos* en sus programas. Algunos de los que han sido aceptados como conocimiento general, cuando estaban trabajando en sus respectivas bases de datos, fueron posteriormente desaprobados. Y así sucesivamente. Existen muchas razones por las que la propiedad de lo *rojo* de la tira —la propiedad causal— no coincidía totalmente con la propiedad de *verdad en inglés*. De este modo, quizá *rojo* pueda ser definido mejor como una expresión relativamente corta en inglés ASCII de algo «creído» como verdad por la caja B (cuyas «creencias» son casi siempre verdad). Esto pudiera satisfacer a algunos, pero otros, quisquillosos, insistirían por varias razones en que esta definición es inexacta, o tiene contraejemplos que no se pueden descartar por ninguna vía que no sea *ad hoc*, y así sucesivamente. Pero, como puntualizan Al y Bo, no podía haber una mejor descripción de la propiedad y ¿no habrían salido ganando los científicos con tal explicación? ¿No ha quedado totalmente resuelto el misterio de las tiras rojas y verdes? Además, ahora que está resuelto, ¿por qué no aceptamos que no hay ninguna esperanza de explicar la regularidad *causal*, con la que hemos comenzado nuestra historia, sin usar *algunos* términos semánticos (o mentalísticos)?

Algunos filósofos argumentan que, aunque la nueva descripción de la regularidad en la actividad del alambre que une las dos cajas pueda ser

utilizada como pretexto para predecir la conducta de la caja B, no se trata, después de todo, de una regularidad *causal*. Verdad y falsedad (o cualquiera de las situaciones estándares que se consideren) son propiedades semánticas y, como tales, enteramente abstractas y, en consecuencia, no pueden causar algo. Un desatino, replicarán otros. Presionar un botón α es la causa de que la luz roja se encienda, del mismo modo que al girar la llave de contacto se pone en marcha nuestro coche. Si lo que ha resultado es que lo que ha sido enviado a través del alambre era simplemente alto voltaje frente a bajo voltaje, o un pulso frente a dos pulsos, todo el mundo estará de acuerdo en que se trata de un sistema causal paradigmático. El hecho de que este sistema resultara ser una máquina de Rube Goldberg no demuestra que la fiabilidad de la conexión entre α y los parpadeos rojos fueran menos causales. Por el contrario, en cada caso específico los científicos pueden trazar la exacta microcausa que explica el resultado[152].

Convencidos con esta línea de razonamiento, otros filósofos comienzan a argumentar que esto demuestra que las propiedades rojo, verde y ámbar no eran *realmente* semánticas o mentalistas sino, tan sólo, imitación de propiedades semánticas, *como* si fueran simplemente propiedades semánticas. Lo que realmente eran el rojo y el verde, eran muy, pero que muy, complicadas propiedades *sintácticas*. Estos filósofos declinaron, no obstante, decir algo más acerca de lo que eran propiedades sintácticas, o explicar cómo incluso los niños podían, rápida y fiablemente, producir ejemplos de estas propiedades o reconocerlas. Sin embargo, los filósofos están convencidos de que allí debería haber una pura descripción sintáctica de la regularidad, dado que, después de todo, los sistemas causales en cuestión eran precisamente ordenadores y los ordenadores son exactamente máquinas sintácticas, incapaces de una real semanticidad.

«Yo supongo», replican Al y Bo, «que si *nos* hubieran encontrado dentro de nuestras cajas negras, jugándoles una treta y siguiendo el mismo esquema, entonces habrían cedido y estarían de acuerdo en que la propiedad operativa causal era una verdad genuina (o una verdad aceptada, en cualquier caso). ¿Pueden presentar alguna buena razón para trazar esta distinción?». Esto conduce a algunos a declarar que en cierto sentido, importante desde luego, Al y Bo *han* estado en las cajas, ya que son responsables de la creación de las respectivas bases de datos como modelos de sus propias creencias. Esto lleva a otros a negar que haya realmente propiedades semánticas o mentalísticas en ningún lugar en el mundo. El contenido, dicen ellos, ha sido *eliminado*. El

debate seguirá durante años, pero el misterio con el cual comenzamos quedó resuelto.

3. *Bloqueadas las salidas*

La historia termina aquí. Sin embargo, la experiencia enseña que no existe experimento mental, por muy claro que se presente, que no pueda ser interpretado erróneamente por un filósofo, así que, con el fin de adelantarme a algunas de las más atractivas interpretaciones falsas, llamaré la atención, de manera poco elegante, sobre algunos de los detalles cruciales y explicaré sus funciones en esta bomba de intuición.

1) Los dispositivos en las cajas A y B no son más que enciclopedias automatizadas; ni siquiera «enciclopedias ambulantes», sino «cajas de verdades». Nada en esta historia presupone o implica que estos dispositivos sean conscientes o *cosas pensantes*, o incluso *agentes*, excepto en el mismo sentido mínimo según el cual un termostato es un agente. Hay sistemas intencionales completamente aburridos, rígidamente ligados al cumplimiento de un objetivo simple y singular. Contienen un gran número de proposiciones verdaderas y la maquinaria inferencial necesaria para generar más verdades y comprobar las que son «verdad», confrontando la proposición candidata con sus bases de datos ya existentes[153].

Dado que los dos sistemas fueron creados de modo independiente, no cabe suponer verosímilmente que contengan exactamente las mismas verdades (real o incluso virtualmente) aunque, para que la broma funcionara tan bien como aseguro en mi historia, debemos suponer un solapamiento muy amplio, por lo que sería altamente improbable que una verdad generada por A no fuera reconocida como tal por B. A mi juicio, dos consideraciones hacen verosímil que esto suceda: a) Al y Bo podrían vivir en países diferentes y tener diferentes lenguajes nativos, pero *habitan el mismo mundo*, y b) mientras que hay miles de millones de proposiciones verdaderas acerca del mundo (nuestro mundo), el hecho de que ambos, Al y Bo, consiguiesen crear bases de datos *útiles* —que contienen la información que es relevante para casi lo más rebuscado de los propósitos humanos— garantizaría un alto grado de solapamiento entre los dos sistemas, que fueron creados independientemente. Aunque Al pudiera saber que a las doce horas del día en que cumplió veinte años su pie izquierdo estaba más cerca del Polo Norte que del Polo Sur, y Bo no hubiera olvidado que su primer profesor de francés se

llamaba Dupont, estas no serían verdades aptas para ser introducidas en su base de datos. Pero, si dudamos de que el mero hecho de que ambos intentaran la creación de una enciclopedia de utilidad internacional, aseguraría una estrecha correspondencia entre las respectivas bases de datos, debemos añadir, desde luego como un detalle poco elegante, el oportuno dato de que Al y Bo, durante sus años como *hackers*, intercambiaron impresiones respecto a los temas que deberían ser cubiertos.

2) ¿Por qué no emparejar a Al con Bob (un compatriota americano) o, en cuanto a eso, por qué no poner simplemente en la caja B un duplicado del sistema de Al? Porque ésta es la esencia (¡uf!) de mi historia: ningún simple emparejamiento *sintáctico,* factible de descubrir, podría explicar la regularidad. Por esta razón el sistema de Bo es el sistema sueco Lisp, para ocultar a los ojos inquisitivos las comunes correspondencias *semánticas* implícitas entre las estructuras de datos consultadas durante la tarea de generar frases en A y la tarea en B de traducir las frases y someterlas a la prueba de la verdad. La idea era crear dos sistemas que exhibieran esta fascinante regularidad de conducta externa, pero que fueran internamente lo más diferentes posible, de modo que tan sólo el hecho de que sus respectivas entrañas fueran sistemáticamente *representaciones de un mundo común* pudiera explicar la regularidad.

3) Podemos hacer una pausa para preguntar si los dos sistemas podrían o no ser tan inescrutables como para ser invulnerables a la ingeniería revertida. La criptografía se ha introducido en estas regiones arcanas de atmósfera enrarecida, por lo que uno debería pensárselo tres veces al menos, antes de dar una opinión. No tengo idea de si alguien puede desarrollar un sólido argumento para demostrar que hay esquemas de cifrado que no pueden romperse, o ¿no es así? Pero, dejando a un lado el tema del cifrado, los *hackers* agradecerían todos los comentarios oportunos y otras señales indicadoras que se colocaran en el «código fuente» cuando se está creando un programa, y que desaparezcan cuando el código fuente es «compilado», dejando detrás un enredo de instrucciones para la máquina, *casi* imposible de descifrar. A veces, es posible en la práctica deshacer la compilación (¿es siempre posible, realmente?) aunque naturalmente no restaurará los comentarios destacará, sin embargo, las estructuras en el lenguaje de más elevado nivel. Mi hipótesis de que los esfuerzos de los científicos al deshacer la compilación del programa y descifrar la base de datos se reducen a la nada puede reforzarse postulando el cifrado, si fuera necesario.

En la historia tal como ha sido contada, podemos estar de acuerdo en que es muy extraño que a los científicos nunca se les haya ocurrido comprobar si existe una traducción al código ASCII de las oleadas de *bits* que recorren el cable entre las dos cajas. ¿Cómo pueden ser tan densas? Bastante razonable: envíese la totalidad del artefacto (cajas A y B y el cable que las conecta) a «Marte» y permítase que los científicos alienígenas intenten reconocer la regularidad. El hecho de que a cause todos los rojos y de que β cause todos los verdes y de que tiras aleatorias de *bits* causen el ámbar, será exactamente tan visible para ellos como para nosotros, pero ellos estarán sin claves con respecto al código ASCII. Para los marcianos, este regalo del espacio exterior exhibirá una regularidad absolutamente misteriosa, más allá de todas las pruebas analíticas, *a menos* que a ellos se les ocurra la idea de que cada caja contiene una *descripción de un mundo*, y que ambas descripciones *son del mismo mundo*. El fundamento de la regularidad se basa en el hecho de que cada caja contenga múltiples relaciones semánticas para las mismas cosas, aunque expresadas en diferente «terminología» y axiomatizadas de forma diferente.

Cuando comenté este experimento mental con Danny Hillis, creador de la máquina de conexión, pensó inmediatamente en una «solución» criptográfica para este rompecabezas y después concedió que mi solución podía aprovecharse, como un caso especial de su solución: «Al y Bo utilizaban el *mundo* como “one-time pad”», una apropiada alusión a la técnica estándar del cifrado. Puede comprobarse esta idea imaginando una variación. Usted y su mejor amigo están a punto de ser capturados por fuerzas hostiles, que pueden conocer el inglés pero no lo suficiente acerca de nuestro mundo. Ambos conocen el código Morse e idean en un *impromptu* el siguiente esquema de cifrado: decir una verdad es una raya, y decir una falsedad es un punto. Vuestros vigilantes os permiten hablar a los dos: «Los pájaros ponen huevos y los sapos vuelan. Chicago es una ciudad y mi pie no está hecho de estaño y el béisbol se juega en agosto», dice usted, respondiendo «No» (raya-punto; raya-raya-raya) a lo que su amigo le haya preguntado. Aunque los vigilantes conocieran el código Morse, a menos que pudieran determinar la verdad y la falsedad de estas sentencias, no podrían detectar las propiedades correspondientes a la raya y al punto. La siguiente variación podría añadirse a nuestra fábula, para darle más sabor: en lugar de embalar los ordenadores en las cajas para enviarlos a Marte, colocaríamos a Al y a Bo dentro de las cajas y los enviaríamos a Marte. Los marcianos quedarían confundidos si Al y Bo aplicaran la jugarreta del código Morse, como con los ordenadores, a menos

que llegasen a la conclusión (obvia para nosotros pero no para los marcianos) de que las cosas que están en las cajas han de ser interpretadas semánticamente.

El corolario de esta fábula es simple. No hay sustituto para la posición intencional; o la adoptamos y explicamos el patrón averiguando el nivel semántico de los hechos, o estaremos siempre desconcertados por la regularidad —la regularidad causal— que allí se manifiesta. La misma moraleja, como hemos visto, puede extraerse de la interpretación de los hechos históricos de la historia de la evolución. Aunque pudiéramos describir, con microdetalles incomparables, todo hecho causal en la historia de cada jirafa que ha vivido alguna vez, a menos que nos elevemos en un nivel o en dos y preguntemos «¿por qué?» —en busca de las *razones* respaldadas por la madre naturaleza—, nunca podremos explicar las regularidades manifiestas, tales como, por ejemplo, el hecho de que las jirafas hayan llegado a tener largos cuellos. A ello se refiere Dewey en la cita recogida previamente en este capítulo.

En esta coyuntura, si uno se comporta como muchos filósofos, nos sentiríamos atraídos por la afirmación de que este experimento mental «funciona» solamente porque las cajas A y B son artefactos cuya intencionalidad, tal como está planteada, es una intencionalidad completamente derivada y artefactual. Las estructuras de datos que hay en sus memorias obtienen su referencia (si consiguen alguna) de la confianza indirecta en los órganos de los sentidos, historias de vidas y propósitos de sus creadores Al y Bo. La fuente real del significado de verdad o semanticidad en los artefactos radica en esos artífices humanos. (Éste era el motivo de la sugerencia de que, en cierto modo, Al y Bo estaban en sus respectivas cajas). Pero yo *podría* haber contado esta historia de manera diferente: dentro de las cajas había dos *robots*, Al y Bo, los cuales habían pasado una gran parte de su vida recorriendo el mundo recolectando datos antes de volver a sus respectivas cajas. Elegí la opción más simple para anticiparme a todas las preguntas acerca de si la caja A o la caja B estaban «realmente pensando», pero ahora podemos volver a considerar estas cuestiones ya que ha llegado finalmente el momento de desechar, de una vez por todas, la sospecha de que la intencionalidad *original* no puede emerger en ninguna «mente» artefactual sin la intervención de un artífice (¿humano?). Supongamos que esto es así. En otras palabras, supongamos que a pesar de las diferencias que pueda haber entre una simple caja de verdades como la caja A y el más hermoso robot imaginable, al ser ambos simplemente artefactos, ninguno de los dos puede

poseer intencionalidad real —u original— sino únicamente la intencionalidad derivada, que emana de su creador. Ahora estamos listos para otro experimento mental, una *reductio ad absurdum* de esta suposición.

4. *Un viaje seguro hacia el futuro*[154]

Supongamos que usted ha decidido, por la razón que sea, que lo que quiere es conseguir experiencia de la vida en el siglo XXV, y supongamos que el único método conocido para mantener vivo su cuerpo, durante todo el tiempo requerido para este viaje, consiste en colocarlo en una especie de dispositivo de hibernación. Supongamos que se trata de una «cámara criogénica» que enfría su cuerpo hasta unos pocos grados por encima del cero absoluto. En esta cámara su cuerpo tendrá la posibilidad de descansar, suspendido en un estado de supercoma, tanto tiempo como usted desee. Puede disponer las cosas para subir a la cámara y a la cápsula que la rodea, dormirse y después, de modo automático, despertarse y salir en el año 2401. Este es, desde luego, un viejo tema de la ciencia ficción.

Diseñar la propia cápsula no es solamente su único problema de ingeniería, dado que la cápsula debe estar protegida y aprovisionada con la energía adecuada (para refrigeración, etc.) durante cuatrocientos años. No podría disponer, naturalmente, de la ayuda de sus hijos y nietos, ya que todos estarán muertos antes del año 2401, y sería poco inteligente presumir que sus descendientes más lejanos, si quedara alguno, tuvieran un verdadero interés por su bienestar. Así que debe diseñar un supersistema para proteger la cápsula y aportar la energía necesaria para cuatrocientos años. En este asunto pueden seguirse dos estrategias básicas. Una sería hacer prospecciones en busca de la localización ideal, la mejor que se pudiera prever, para una instalación fija que estuviera bien aprovisionada con agua, luz solar y cualquier otra cosa que su cápsula (y el propio supersistema) necesitase durante el tiempo de la experiencia. El principal obstáculo para esta instalación o «planta» es que no puede trasladarse si acecha un peligro: por ejemplo, si alguien decide construir una autopista precisamente donde estuviera localizada la cápsula. La estrategia alternativa es bastante más compleja y costosa, pero evitaría este inconveniente: diseñar un dispositivo móvil para albergar su cápsula, junto con los sensores necesarios y los dispositivos de alarma, de modo que pudiera moverse cuando su posición corriera peligro, y buscar nuevas fuentes de energía y materiales de

reparación, si fuera necesario. En resumen, construir un robot gigante e instalar en su interior la cápsula (con el viajero hacia el futuro dentro).

Estas dos estrategias básicas están copiadas de la naturaleza: corresponden aproximadamente a la división entre plantas y animales. Dado que la segunda estrategia, más compleja, se adapta mejor a nuestros propósitos, podemos suponer que se decide por construir un robot para albergar la cápsula. Debe intentar un diseño de este robot que, por encima de todo, «escoja», como es natural, acciones diseñadas en favor de sus intereses como viajero. No denomine puntos de «elección» a los simples interruptores colocados en el sistema de control del robot, si piensa que esto implicaría que el robot posea libre albedrío o consciencia, ya que no deseo introducir este tipo de contrabando en el experimento mental. Mi proposición es incontrovertible: el poder de cualquier programa de un ordenador radica en su capacidad para ejecutar instrucciones *ramificadas*, optando por una rama u otra, según algunas pruebas que lleva a cabo con los datos disponibles en ese momento, y mi proposición es que cuando planifique el sistema de control del robot, debe ser lo bastante inteligente para intentar estructurarlo de modo que en cualquier oportunidad de ramificación ante las que se encuentre, tenderá a seguir aquella rama con la más elevada probabilidad de servir a *sus propios intereses*. Esta es, después de todo, la *raison d'être* del artefacto en su conjunto. La idea de diseñar un *hardware* y un *software* que se adapten muy ajustadamente a los intereses de un ser humano en particular no es ciencia ficción, aunque los problemas concretos de diseño a los que se enfrentan los constructores de robots serían unos retos de ingeniería muy difíciles, algo más allá del estado actual de la ingeniería. Esta entidad móvil necesitaría un sistema de «visión» para guiar su locomoción y de otros sistemas «sensores», además de las capacidades de automonitorización que le informen de sus necesidades.

Teniendo en cuenta que como viajero hacia el futuro se encontrará en estado de coma a través de todo el recorrido y, en consecuencia, no podrá mantenerse en estado de alerta para guiar y planificar las estrategias, tendrá que diseñar el supersistema del robot que genere sus propios planes, en respuesta a las cambiantes circunstancias a lo largo de los siglos. Debe «saber» cómo «buscar» y «reconocer» y, posteriormente, explotar los recursos energéticos, cómo moverse por el territorio más seguro, cómo «anticiparse» y, de este modo, evitar los peligros. Con tanto por hacer, y hacerlo rápido, es mejor que dependa, cuanto pueda, de la economía: proporcionar a su robot no más habilidad discriminatoria de la que

probablemente necesita para distinguir las diferentes necesidades presentes en el mundo, dada su particular constitución.

Su tarea como presunto viajero se hará más difícil si no puede contar con que su robot sea el único robot en los alrededores dedicado a esta misión. Supongamos que, además de la gente y otros animales que hayan aparecido durante siglos, haya otros robots, muchos robots diferentes (y quizá «plantas» también) compitiendo con su robot por la energía y la seguridad. (¿Por qué podría algo así ponerse de moda? Supongamos que se consigue por adelantado la irrefutable evidencia de que viajeros de otras galaxias llegarán a nuestro planeta en el 2.401. Yo, por lo menos, estaría ansioso por estar allí, para encontrarme con ellos, y si conservarme congelado fuera mi única posibilidad, estaría tentado de seguir este camino). Si ha de planificar el encuentro con otros agentes robóticos que actúen en nombre de otros clientes, debe ser inteligente para diseñar su robot con la suficiente sofisticación de su sistema de control que le permita calcular los riesgos y los beneficios de cooperar con otros robots o de constituir alianzas en su propio beneficio. Sería muy poco inteligente de su parte si supusiera que otros clientes seguirán la regla del «vive y deja vivir»; puede haber robots baratos que funcionen como «parásitos», por ejemplo, a la espera de penetrar en su costosa construcción y explotarla. Cualquier cálculo que su robot haga acerca de estas amenazas y oportunidades ha de ser «rápido y sin contemplaciones»: no hay un método infalible para diferenciar amigos de enemigos, o traidores de cumplidores de promesas, de modo que tendrá que diseñar el robot para que sea, como un jugador de ajedrez, alguien que toma decisiones y acepta riesgos como respuesta a la presión del tiempo.

El resultado de este proyecto de diseño sería un robot capaz de demostrar un alto grado de autocontrol. Dado que debe traspasar al robot el control de pequeños detalles, y en tiempo real, una vez que esté dormido, será tan «remoto» como lo eran los ingenieros de Houston cuando le otorgaron su autonomía a la nave espacial Viking (véase el capítulo 12). Como agente autónomo, el robot será capaz de derivar sus propios objetivos subsidiarios de la evaluación de su estado actual y la importancia de tal estado para su objetivo último (preservarle a usted hasta el año 2401). Estos objetivos subsidiarios, que responderán a las circunstancias que no puede predecir en detalle (si pudiera debería preparar las mejores respuestas ante ellas), pueden llevar al robot bastante lejos en proyectos como éste que duran siglos, algunos de los cuales pueden ser poco aconsejables a pesar de sus mejores esfuerzos. Su robot puede embarcarse en acciones antitéticas para su propósito de

alcanzar el futuro, incluso en acciones suicidas, si otro robot le convence de subordinar la propia misión de su vida a alguna otra.

El robot que hemos imaginado estaría muy absorbido en su mundo y en sus proyectos, siempre dirigidos, en último término, por todo aquello que reste del objetivo que usted estableció en el momento de entrar en la cápsula. Todas las preferencias del robot serán como descendencia de las preferencias que inicialmente le confirió, con la esperanza de que le conduciría al siglo XXV, aunque esto no garantiza que las acciones tomadas a la luz de las preferencias descendientes del robot continúen respondiendo, de modo directo, a sus mejores intereses como viajero hacia el futuro. Desde su egoísta punto de vista, esta es su esperanza, pero los proyectos del robot están fuera de su control directo hasta que despierte. El robot tendrá alguna representación interna de sus objetivos más importantes en un momento dado, de su *summum bonum*, pero si ha caído entre compañeros robóticos persuasivos, de la calaña de los que hemos imaginado, la mano de hierro de la ingeniería que inicialmente lo diseñó estará comprometida. Seguirá siendo un artefacto, actuando solamente como su ingeniería le permite actuar, pero siguiendo un conjunto de *desiderata*, en parte de su diseño.

Y aun así, de acuerdo con la hipótesis que decidimos explorar, este robot no exhibirá otra cosa que no sea intencionalidad *derivada*, dado que es precisamente un artefacto, creado para servir sus intereses como viajero al futuro. Podemos denominar a esta posición «cliente-centrismo» con respecto al robot: yo soy la fuente original de todo el significado derivado dentro de mi robot, no importa lo lejos que se haya apartado de su objetivo inicial. El robot es exactamente una máquina de supervivencia, diseñada para conducirme con seguridad hacia el futuro. El hecho de que ahora se encuentre implicado arduamente en proyectos que sólo de manera remota están conectados con mis intereses, y que incluso pueden ser antitéticos con ellos, no califica, de acuerdo con nuestra hipótesis, como una genuina intencionalidad a ninguno de sus estados de control, o a sus estados «sensorial» y «perceptivo». Si queremos profundizar en esta idea del cliente-centrismo, entonces debemos estar dispuestos a extraer la consiguiente conclusión de que nunca disfrutamos de estado alguno con intencionalidad *original*, dado que somos una máquina de supervivencia diseñada originalmente con el propósito de preservar nuestros genes hasta que puedan replicarse. Después de todo, nuestra intencionalidad se deriva de la intencionalidad de nuestros genes egoístas. ¡*Ellos* son la fuente de toda intencionalidad, no nosotros!

Si esta situación no le atrae, considere la posibilidad del extremo opuesto. Admita que un artefacto bastante fantástico —algo parecido a los imaginados robots— *pueda* exhibir intencionalidad real, dada su rica implicación funcional en el medio ambiente y su valentía en la autoprotección y el autocontrol[155]. Este robot, como nosotros, debe su existencia a un proyecto cuyo objetivo era crear una máquina de supervivencia, pero, como nosotros, ha adquirido una cierta autonomía, ha llegado a ser un *locus* de autocontrol y de autodeterminación, no por milagro, sino justamente afrontando problemas a lo largo de su propia vida y, más o menos, resolviéndolos; problemas de supervivencia que le ha planteado el mundo. ¿Máquinas de supervivencia más simples —las plantas, por ejemplo— nunca consiguen los altos niveles de autodefinición que hace posible la complejidad de su robot; si las consideramos *solamente* como máquinas de supervivencia para sus comatosos habitantes, entonces no encontraremos patrones inexplicados en *sus* conductas.

Si prosigue por este camino, que naturalmente yo recomiendo, entonces debe abandonar la objeción de Searle y de Fodor a la inteligencia artificial, «basada en fuertes principios». El imaginado robot, aunque sea una hazaña de ingeniería difícil e improbable, no es una imposibilidad y nadie lo afirma. Searle y Fodor aceptan la posibilidad de este robot, aunque discuten su «*status* metafísico»; a pesar de la destreza con que los robots solucionan sus problemas, sostienen que su intencionalidad no sería real. Esto es hilar demasiado fino. Yo recomiendo descartar este repudio sin esperanza y reconocer que el significado que este robot descubriría en su mundo, y que aprovecharía en su comunicación con otros, sería exactamente tan real como el significado del que nosotros disfrutamos. En este sentido, nuestros genes egoístas pueden considerarse como la fuente original de nuestra intencionalidad y, en consecuencia, de todo significado que podamos contemplar o conjurar, aun cuando podamos trascender nuestros genes, usando nuestra experiencia y, en particular, la cultura en la que estamos embebidos, para construir un *locus* de significado casi enteramente independiente (o «trascendente») sobre la base que nuestros genes han aportado.

En mi opinión esta es una solución enteramente compatible —y, además, inspiradora— que resuelve la tensión existente entre el hecho de que yo, como persona, me considere a mí mismo una fuente de significación, un árbitro de lo que sucede y por qué sucede, y el hecho de que, al mismo tiempo, sea un miembro de la especie *Homo sapiens*, un producto no

milagroso de varios miles de millones de años de I+D, que no disfruta de característica alguna que no haya surgido del mismo conjunto de procesos, por una u otra vía. Yo sé que otros encuentran esta visión tan chocante que retornan con renovada vehemencia a la convicción de que en algún lugar, de alguna manera, *ha* de haber un bloqueo contra el darvinismo y la inteligencia artificial. He intentado demostrar que la peligrosa idea de Darwin lleva consigo la implicación de que no hay tal bloqueo. De la verdad del darvinismo se sigue que todos nosotros somos artefactos de la madre naturaleza, pero nuestra intencionalidad no es menos real por ser un efecto de millones de años de I+D insensato y algorítmico, en lugar de un regalo que procede de lo alto. Jerry Fodor puede hacer chistes con la ridícula idea de nuestro ser como artefacto de la madre naturaleza, pero la risa suena hueca: la única visión alternativa postula un tipo u otro de ganchos celestes. El impacto que produce esta conclusión puede bastar para hacer más simpáticos los desesperanzados intentos de Chomsky o de Searle por ocultar la mente tras un velo de misterio impenetrable, o los estériles intentos de Gould por escapar de la implicación de la selección natural: una serie algorítmica de grúas operando sobre las más elevadas formas de diseño.

O puede que esta conclusión nos incite a buscar un salvador en otro lugar. ¿Lo será el matemático Kurt Gödel al probar un gran teorema que demuestra la imposibilidad de la inteligencia artificial? Muchos han pensado así y recientemente sus presentimientos han recibido un fuerte impulso de parte de Roger Penrose, uno de los físicos y matemáticos más eminentes, en su libro *La nueva mente del emperador*, al cual dedicaré el siguiente capítulo.

CAPÍTULO 15
La nueva mente del emperador y otras fábulas

I. *La espada en la piedra*

> En otras palabras, si se espera que una máquina sea infalible, entonces no podrá ser inteligente. Hay varios teoremas que dicen casi exactamente lo mismo. Pero estos teoremas no dicen nada acerca de cuánta inteligencia puede manifestarse si una máquina no pretende ser infalible.
>
> Alan Turing,
> *ACF Reports of 1946 and Other Papers*

A lo largo de los años, los intentos de utilizar el teorema de Gödel para probar algo importante acerca de la naturaleza de la mente humana han estado rodeados de una inaprensible atmósfera de aventura. Hay algo extrañamente emocionante en la posibilidad de «utilizar la ciencia» con este objeto. Creo que estoy en condiciones de participar en este asunto. El texto clave no es el de Hans Christian Andersen sobre los nuevos trajes del emperador, sino la leyenda del rey Arturo, en la que se relata la historia de la espada clavada en la piedra. Alguien (nuestro héroe, desde luego) posee una virtud especial, casi mágica, que no se manifiesta en la mayoría de las circunstancias, aunque puede ser revelada, de manera inequívoca, en circunstancias especiales: aquel que consiga extraer la espada de la piedra en la que se encuentra clavada posee esa virtud; si no puede extraerla, no la posee. Esto es una hazaña o un fracaso a la vista de todos; no requiere ninguna interpretación especial ni una específica declaración personal. Si saca la espada de la piedra ha ganado, es todo lo que hay que hacer.

Lo que el teorema de Gödel promete al que tiene inclinaciones románticas es una prueba dramática similar acerca de lo «especial» que es la mente humana. El teorema de Gödel define una habilidad, según parece, que una mente humana germina puede realizar pero que ningún impostor, ningún simple robot controlado por algoritmos podría realizar. Los detalles técnicos de la prueba de Gödel, en sí misma, no nos conciernen; ningún matemático duda de sus fundamentos. Toda la controversia estriba en *cómo utilizar* el

poder del teorema para probar algo acerca de la naturaleza de la mente. La debilidad en un argumento de este tipo se presenta en el crucial paso empírico: el paso donde nosotros *observamos* a nuestros héroes (nosotros mismos, nuestros matemáticos) haciendo algo que el robot simplemente no puede hacer. ¿Es esta hazaña comparable a arrancar la espada de la piedra, una hazaña que no tiene similitudes verosímiles, o se trata de una hazaña que no puede ser distinguida fiablemente de las simples aproximaciones? Esta es la cuestión crucial, y ha habido gran confusión sobre lo que representa esta distinguida hazaña. En parte la confusión puede ser culpa del propio Kurt Gödel, ya que él pensó que había demostrado que la mente humana *debe* ser un gancho celeste.

En 1931, Gödel, un joven matemático de la Universidad de Viena, publicó su prueba, uno de los más importantes y sorprendentes resultados matemáticos del siglo XX, estableciendo un límite absoluto sobre la prueba matemática, que es, en realidad, completamente chocante. Recordemos la geometría euclidiana que estudiamos en el bachillerato, en la cual aprendimos a crear pruebas formales de los teoremas de la geometría, a partir de una lista básica de axiomas y definiciones, usando una lista fija de reglas de inferencia. Lo aprendimos a través de una *axiomatización* de la geometría plana. Recordemos cómo el profesor dibujaba un diagrama geométrico en la pizarra, nos mostraba un triángulo con varias líneas rectas que interseccionaban sus lados de varias formas, reuniéndose en varios ángulos y, luego, nos formulaba preguntas como: ¿Se producirá la intersección de estas dos líneas en ángulo recto?, ¿es este triángulo que está aquí congruente con el triángulo que está allí? A menudo la respuesta era obvia: podíamos *ver* con exactitud que la intersección de las líneas sería en ángulo recto y que los triángulos eran congruentes. Pero otro asunto era —de hecho, un trabajo duro y tedioso aunque estimulante— probarlo a partir de los axiomas, formalmente, de acuerdo con reglas estrictas. Cuando el profesor dibujaba un nuevo diagrama en la pizarra, ¿llegamos a sospechar que se trataba de hechos de la geometría plana que podíamos *ver* que eran verdad, pero que no podíamos probar, ni siguiera trabajando un millón de años?, ¿nos parecía obvio que al ser incapaces de diseñar la prueba de alguna posible verdad geométrica, lo considerásemos como un fracaso personal? Quizá pensamos: «¡*Debe haber* una prueba, dado que es *verdad*, aunque yo nunca logre encontrarla!».

Ésta es una opinión muy plausible, pero lo que Gödel probó, más allá de cualquier duda, es que cuando se trata de axiomatizar la simple *aritmética* (no la geometría plana) hay verdades que nosotros podemos ver que son

verdaderas, pero cuya verdad *nunca* podremos demostrar. En realidad, esta afirmación debe ser cuidadosamente delimitada y defendida: para cualquier sistema axiomático *específico* que sea *coherente* (no sutilmente autocontradictorio, lo que sería un fallo descalificante) debe haber una proposición de aritmética, ahora conocida como la proposición de Gödel de este sistema, que no se puede probar dentro del sistema, pero que es verdad. (De hecho, debe haber muchas de estas proposiciones verdaderas, pero sólo necesitamos una para la cuestión que ahora nos interesa). Podemos cambiar sistemas, y probar esta primera proposición de Gödel en el siguiente sistema axiomático que hemos escogido, pero consecutivamente producirá su propia proposición de Gödel, si es consistente, y así en adelante. Ninguna axiomatización simple y consistente de la aritmética puede probar todas las verdades de la aritmética.

Todo esto nos puede parecer poco interesante, dado que rara vez deseamos *probar* hechos aritméticos; damos por probada la aritmética sin prueba. Pero es posible diseñar sistemas axiomáticos de tipo euclidiano para la aritmética —por ejemplo, los axiomas de Peano, que cubren el concepto lógico de los números naturales— y probar verdades tan simples como «2 + 2 = 4», verdades de nivel medio tan obvias como «los números divisibles por 10 son también divisibles por 2», y verdades no obvias como «no existe el número primo más grande». Antes de que Gödel diseñara su prueba, el objetivo de derivar todas las verdades matemáticas de un simple conjunto de axiomas fue considerado por los matemáticos y los lógicos como el gran proyecto, difícil pero alcanzable, algo así como el viaje a la Luna o el proyecto del genoma humano para los matemáticos de aquellos días. Pero ahora sabemos que es absolutamente imposible cumplir ese proyecto. Esto es lo que el teorema de Gödel establece.

Ahora bien, ¿qué tiene que ver todo esto con la inteligencia artificial o con la evolución? Gödel demostró su teorema unos años antes de la invención del ordenador electrónico, aunque entonces Alan Turing llegó y extendió la implicación de este teorema abstracto, demostrando, en efecto, que cualquier procedimiento formal de prueba, del tipo propuesto por el teorema de Gödel, es equivalente a un programa de ordenador. Gödel había diseñado un método para colocar *todos los posibles sistemas axiomáticos* en «orden alfabético». De hecho, todos pueden ser alineados en la Biblioteca de Babel, y Turing demostró que este grupo era un subgrupo de otro grupo en la Biblioteca de Babel: un grupo con *todos los posibles ordenadores*. No importa de qué material esté hecho un ordenador; lo que importa es el algoritmo que allí

opera; y dado que todo algoritmo es especificable de modo finito, es posible diseñar un lenguaje uniforme para describir únicamente cada algoritmo y colocar todas las especificaciones en orden alfabético. Turing diseñó precisamente ese sistema en el cual todo ordenador —desde su portátil a los más grandes superordenadores en paralelo que aún puedan construirse— tiene una única descripción como lo que llamamos una *máquina de Turing*. A cada una de las máquinas de Turing se les da un único número de identificación, su número en la Biblioteca de Babel, si les parece. El teorema de Gödel puede entonces reinterpretarse al decir que cada una de aquellas máquinas de Turing que resultan ser un algoritmo coherente para probar las verdades de la aritmética (y no sorprende que estas sean un «vasto» pero «evanescente» subgrupo de todas las posibles máquinas de Turing) lleva asociada en *ella* una proposición de Gödel, es decir, una verdad de la aritmética que no puede probarse. De modo que esto es lo que nos dice Gödel, apoyado por Turing en el mundo de los ordenadores: todo ordenador que sea un demostrador de las verdades coherentes de la aritmética tiene un talón de Aquiles, una verdad que nunca puede probar, aunque siga funcionando hasta el día del Juicio Final. ¿Y por qué?

El propio Gödel pensó que la implicación de su teorema era que los seres humanos —al menos los matemáticos entre nosotros— no pueden ser máquinas, debido a que pueden hacer cosas que ninguna máquina podría hacer. Más exactamente, al menos alguna parte de un ser humano no puede ser una mera máquina, ni siquiera una enorme colección de artefactos. Si los corazones son máquinas bombeadoras y los pulmones, máquinas que intercambian el aire y los cerebros, máquinas computadoras, las mentes de los matemáticos, pensó Gödel, no pueden ser sus cerebros, dado que las mentes de los matemáticos pueden hacer algo más de lo que una máquina computadora puede hacer.

¿Qué es, exactamente, lo que puede hacer? Este es el problema de definir la hazaña para la gran prueba empírica. Es una tentación pensar que ya hemos visto un ejemplo: las mentes pueden hacer lo que solíamos hacer cuando mirábamos a la pizarra en la clase de geometría; utilizando algo así como «intuición» o «juicio» o «pura comprensión», pueden *justamente ver* que ciertas proposiciones aritméticas son verdaderas. La idea sería que no tienen necesidad de depender de complicados algoritmos para generar su conocimiento matemático, dado que tienen un talento especial para captar las verdades matemáticas que trascienden completamente a los procesos algorítmicos. Recordemos que un algoritmo es una receta que puede ser

seguida por serviles ignorantes, o incluso máquinas; no es necesaria comprensión alguna. Los matemáticos inteligentes parecen, por contraste, capaces de *utilizar* su comprensión para ir más allá de lo que estos mecánicos ignorantes pueden hacer. Pero, aunque esto parece ser lo que pensó el propio Gödel, y ciertamente viene a expresar lo que popularmente se entiende por teorema de Gödel, parece más difícil de demostrar de lo que parece a primera vista. ¿Cómo podríamos, por ejemplo, distinguir el caso de alguien (o de algo) que «capta la verdad» de una proposición matemática, del caso de alguien (o de algo) que realiza correctamente conjeturas a la buena de Dios? Podemos entrenar a un loro para que diga «verdadero» y «falso» cuando se escriban varios símbolos en una pizarra; ¿cuántas respuestas sin error tendría que dar el loro para que estuviera justificado el que creyéramos que el loro, después de todo, tiene una mente inmaterial (o quizá sea un matemático humano disfrazado de loro) (Hoístadter 1979)?

Éste es el problema que siempre ha creado conflictos a aquellos que desean utilizar el teorema de Gödel para probar que nuestras mentes son ganchos celestes y no precisamente viejas y aburridas grúas. No pretendo decir que los matemáticos, al contrario que las máquinas, puedan probar cualquier verdad aritmética si lo que entendemos por «probar» es lo que Gödel entiende por «probar» en su prueba, donde demuestra que los seres humanos —o los ángeles, si los hubiera— no pueden hacerlo (Dennett 1970); *no hay* prueba formal de una proposición del sistema de Gödel dentro del sistema. Un famoso y precoz intento de utilizar el poder del teorema de Gödel fue el del filósofo J. R. Lucas [1961, 1970], quien decidió definir así la crucial hazaña: «presentar como verdadera» una cierta proposición, de Gödel o de cualquier otro. Sin embargo, esta definición se enfrenta con insolubles problemas de interpretación, arruinando el carácter definitivo (expresado con la imagen de «la espada en la piedra») de la vertiente empírica del argumento (Dennett 1970, 1972; véase también Hofstadter 1979). Podemos ver ahora más claramente cuál es el problema considerando varias hazañas relacionadas, reales e imaginarias.

René Descartes, en 1637, se preguntó cómo se podría distinguir un genuino ser humano de una máquina y llegó a «dos medios muy seguros»:

El primer medio es que ellas [las máquinas] nunca pueden usar palabras, o poner juntos otros signos, como hacemos nosotros para expresar nuestros pensamientos. Para que podamos ciertamente concebir una máquina así construida que pronuncie palabras e incluso que pronuncie palabras que correspondan a acciones corporales que causen un cambio en sus órganos (por ejemplo, si se toca a alguien en un punto, pregunta qué desea, si se toca en otro, llora y dice que le esta haciendo daño, y así sucesivamente). Pero no es concebible que tal máquina deba producir diferentes disposiciones de palabras para dar una

respuesta con significado a lo que se dice en su presencia, como puede hacer el más estúpido de los hombres. El segundo es que incluso en el caso de que tales máquinas puedan hacer algunas cosas tan bien como nosotros las hacemos, o quizá mejor, inevitablemente fallarían en otras, lo cual revelaría que estaban actuando no a través de la comprensión, sino solamente por la disposición de sus órganos. Mientras que la razón es un instrumento universal que puede ser usado en roda clase de situaciones, estos órganos necesitan alguna particular disposición, para cada acción concreta; en consecuencia es, a todos los propósitos prácticos, imposible para una máquina tener suficientes órganos diferentes para actuar en todas las contingencias de la vida, del mismo modo en que nuestra razón actúa (*Discurso del método*).

En 1950, Alan Turing se hizo la misma pregunta y llegó a la misma prueba —algo más rigurosamente descrita—, a la que llamó el *juego de la imitación* y que denominaremos la prueba de Turing. Colocó a los dos competidores —uno humano y otro un ordenador— dentro de unas cajas (realmente) y mantuvo conversaciones con cada uno de ellos; si el ordenador puede convencerle de que es el ser humano, él gana el juego de la imitación. Sin embargo, el veredicto de Turing fue bastante diferente del de Descartes:

Creo que en cincuenta años, aproximadamente, serán posibles programas de ordenadores con una capacidad de almacenamiento de alrededor de 10^9, para hacerles jugar tan bien el juego de la imitación que un interrogador medio no tendrá más que un 70 por ciento de oportunidades de hacer la identificación correcta después de cinco minutos de preguntas. La cuestión original, «¿Pueden las máquinas pensar?», creo que no tiene sentido y por lo tanto no merece ser discutida. Sin embargo, creo que al finalizar este siglo el uso de palabras y la educación de la opinión general se habrá modificado lo suficiente como para que seamos capaces de hablar de máquinas pensantes sin que nos contradigan (Turing 1950:435).

La última profecía de Turing ha demostrado ser correcta: «el uso de palabras y la educación de la opinión general» han sido ya «bastante modificados», de modo que se puede hablar de máquinas pensantes sin esperar ser contradecido acerca de los principios generales. Para Descartes era inconcebible la noción de máquina pensante, e incluso si, como muchos creen hoy, ninguna máquina pasará jamás la prueba de Turing, casi nadie sostendría actualmente que la idea en sí es inconcebible.

Quizás esta marea de cambio en la opinión pública ha sido apoyada paralelamente por el progreso de los ordenadores en otras «hazañas», tales como jugar a las damas y al ajedrez. En una conferencia pronunciada el año 1957, Herbert Simon (Simon y Newell 1958) predijo que un ordenador sería el campeón del mundo de ajedrez en menos de una década, un caso clásico de hiperoptimismo, como así ha resultado. Varios años después el filósofo Hubert Dreyfus [1965] predijo que ningún ordenador sería nunca capaz de jugar bien al ajedrez dado que jugar al ajedrez requiere «intuición», pero el propio Dreyfus pronto fue derrotado al ajedrez por un programa de ordenador,

circunstancia que provocó bastante alegría entre los investigadores de la IA. Al programa de damas de Art Samuel le han seguido literalmente cientos de programas de ajedrez que ahora compiten en torneos, tanto con seres humanos como con otros programas, y probablemente pronto serán capaces de batir a los mejores jugadores de ajedrez del mundo.

Pero ¿es el juego del ajedrez una prueba apropiada como lo fue la de «arrancar la espada que estaba clavada en la piedra»? Dreyfus pudo haberlo pensado alguna vez, ya que a un distinguido predecesor —Edgar Allan Poe— su certeza sobre esta cuestión le llevó a desenmascarar uno de los grandes fraudes del siglo XIX, el «automaton» de ajedrez de Von Kempelen. En el siglo XVIII, el gran Vaucanson había construido maravillas mecánicas que fascinaron a la nobleza y a otros clientes provechosos, mediante la exhibición de artilugios que incluso hoy nos inspiran escepticismo. ¿Pudo realmente el pato con mecanismo de relojería de Vaucanson hacer lo que se dice que hizo? «Cuando se le arrojó maíz al suelo, el pato estiró el cuello para cogerlo, tragarlo y digerirlo» (Poe, «El jugador de ajedrez de Maelzel»). Otros ingeniosos artífices y tramposos siguieron la ola de Vaucanson y desarrollaron la artesanía de los simulacros mecánicos hasta tal punto que, en 1769, uno de estos artífices, el barón von Kempelen, pudo explotar la fascinación pública con tales artilugios creando un deliberado engaño: un autómata que *presuntamente* podía jugar al ajedrez.

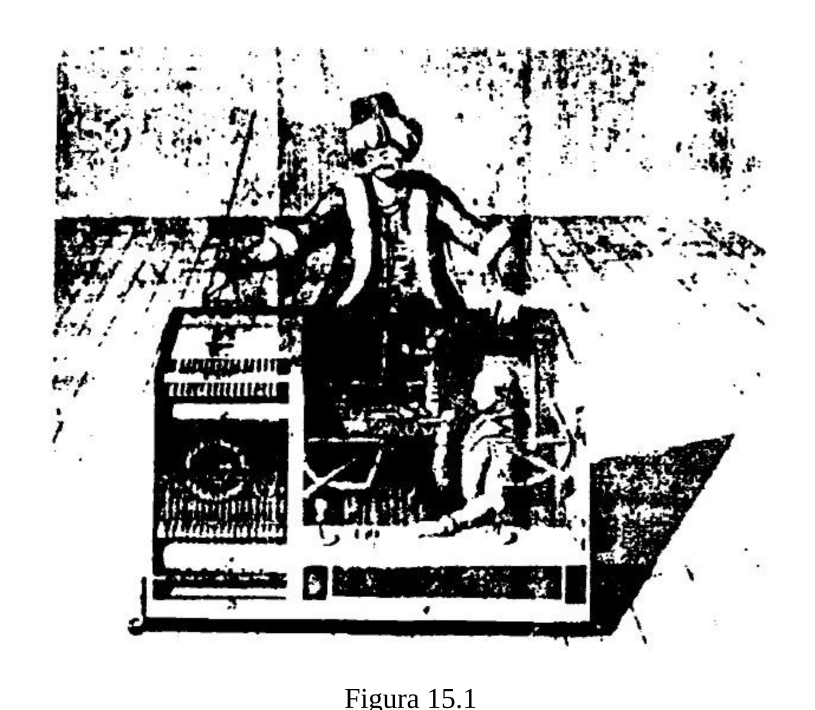


Figura 15.1

El «automaton» que jugaba al ajedrez de Von Kempelen.

La máquina original de Von Kempelen pasó a manos de Johann Nepomuk Maelzel[156], quien hizo algunas mejoras y revisiones y, poco después, provocó una conmoción en los comienzos del siglo XIX al exhibir la máquina de ajedrez de Maelzel, sin garantizar completamente que fuera una máquina, ambientando toda la representación del espectáculo (por el que cobraba un penique) con muchas de las ostentaciones habituales de los magos, lo que despertó sospechas en todo el mundo. La figura de un turco, claramente mecánica, estaba sentada ante una mesa de escritorio, sospechosamente cerrada, cuyas puertas y cajones se abrían secuencialmente (aunque nunca todos a la vez), lo que permitía al público «ver» que no había nada dentro, excepto la maquinaria. A partir de ese momento, la figura vestida como un otomano comenzaba a jugar una partida de ajedrez, cogiendo y moviendo las piezas sobre un tablero en respuesta a los movimientos de su oponente humano y ¡casi siempre ganaba! Pero ¿había allí dentro, literalmente, un homúnculo, un hombre pequeño que hacía el trabajo mental? Si la IA es posible, el escritorio *debería* estar repleto de un montón de poleas y otras piezas de maquinaria. Si la IA es imposible, entonces debería haber un gancho celeste dentro del escritorio, una mente haciéndose pasar por una máquina.

Edgar Allan Poe estaba absolutamente seguro de que la máquina de Maelzel ocultaba un ser humano y su ingeniosa investigación confirmó sus sospechas, tal como él describe en detalle, con un apropiado aire de triunfo, en un artículo publicado en el *Southern Literary Messenger* (1836). Tan interesante como su razonamiento acerca de cómo se perpetró el engaño, es su reflexión —en una carta que acompaña la publicación de este artículo— sobre por qué tenía que tratarse de un engaño, una línea de argumentación que puede ser, perfectamente, el eco de la prueba de John Locke (véase el capítulo 1).

Nosotros nunca hemos dado nuestro asentimiento, en ningún momento, a la opinión prevalente de que el señor Maelzel no utiliza una autoría humana. La opinión de que tal autoría humana es empleada no puede ser cuestionada, a menos que se demuestre a plena satisfacción que el hombre es capaz de impartir intelecto a la materia: una *mente* es tan imprescindible en las operaciones del juego de ajedrez como lo es en la prosecución de una cadena de razonamientos abstractos. Recomendamos a aquellos cuya credulidad ha sido en este caso cautivada por una apariencia plausible y a todos, crédulos o no, quienes admiren una ingeniosa serie de razonamientos inductivos, que lean este artículo atentamente; todos y cada uno, tras su cuidadosa lectura, deben quedar convencidos de que una simple máquina no puede comprometer al intelecto con estas intrincadas exigencias del juego… (Poe 1836b:89).

Ahora sabemos que, a pesar de lo convincente que solía resultar este razonamiento, su fundamento ha sido destruido por Darwin y la particular conclusión que Poe extrae acerca del ajedrez ha sido definitivamente refutada por la generación de artífices que ha seguido los pasos de Art Samuel. ¿Qué pasa entonces con la prueba de Descartes, ahora conocida como prueba de Turing? Esta prueba ha generado controversia, desde que Turing propuso su versión tan bellamente operativa, y llevó a cabo una serie de competiciones reales, aunque restringidas, que confirman lo que cualquiera que hubiera reflexionado cuidadosamente sobre la prueba de Turing ya sabía (Dennett 1985): es vergonzosamente fácil engañar a los jueces ingenuos y astronómicamente difícil engañar a los jueces expertos; el problema, una vez más, es no contar como prueba con una hazaña apropiada, como la de «arrancar la espada de la piedra», con lo cual el asunto quedaría resuelto. Mantener una conversación o ganar una partida de ajedrez no es una hazaña apropiada, la primera debido a que es demasiado abierta para que un concursante se asegure una victoria ambigua, a pesar de su gran dificultad, y la segunda porque es demostrable, después de todo, dentro de las posibilidades de una máquina. ¿Pueden las implicaciones del teorema de Gödel facilitar una competición mejor? Supongamos que hemos colocado a un matemático en la caja A y a un ordenador —del tipo que se quiera— en la caja B, y planteamos a cada uno de ellos cuestiones acerca de la verdad o la

falsedad de proposiciones aritméticas. ¿Sería *ésta* una prueba capaz de desenmascarar a la máquina? El problema es que todos los matemáticos humanos cometen errores y el teorema de Gödel no ofrece un veredicto acerca de la probabilidad, sino sólo sobre la imposibilidad de una detección por un algoritmo de una verdad menos que perfecta. No parece, entonces, que haya ninguna prueba aritmética apropiada que podamos proponer a las cajas A y B para distinguir, con claridad, al hombre de la máquina.

Esta dificultad ha sido considerada, en general, como un bloqueo sistemático de cualquier argumento derivado del teorema de Gödel acerca de la imposibilidad de la inteligencia artificial. Es cierto que todos los que trabajan en JA conocen el teorema de Gödel y todos ellos han continuado con su trabajo sin preocuparse. De hecho, el clásico libro de Hofstadter *Gödel Escher Bach* (1979) puede leerse como la demostración de que Gödel es un reacio *defensor* de la IA, que aporta puntos de vista esenciales sobre las vías a seguir hacia una IA «fuerte», sin mostrar la futilidad de este campo de investigación. Pero Roger Penrose, profesor de matemáticas en Oxford, en la cátedra Rouse Ball, piensa de otra manera. Su reto ha de ser tomado con seriedad, incluso por los que están convencidos, como yo y otros dedicados a la IA, de que Penrose está cometiendo un error muy simple. Cuando apareció el libro de Penrose señalé, en una reseña, dónde estaba el problema: su argumento es excesivamente barroco, y erizado con detalles de física y de matemática,

> y es improbable que una iniciativa de este tipo sucumbiera a un simple y aplastante golpe de vista lanzado por su creador; que el argumento pudiera ser «refutado» por una simple observación. Así que yo me resisto a dar crédito a mi observación de que Penrose parece cometer un error elemental precisamente al comienzo y, en cualquier caso, falla en darse cuenta o rechazar lo que parece ser una objeción obvia (Dennett 1989b).

Mi sorpresa e incredulidad halló pronto eco, en primer lugar por la habitual serie de comentaristas que provoca un artículo de Penrose (basado en su libro), publicado en la revista *Behavioral and Brain Sciences*, con la finalidad específica de abrir una discusión escrita, y en segundo lugar por la respuesta de Penrose cuando le llegó su turno. En el texto titulado «The Nonalgorithmic Mind» (1960), al replicar a sus críticos, Penrose expresa una moderada sorpresa ante la dureza del lenguaje utilizado por algunos: ¡«completamente falaz», «erróneo», «fallo letal» e «inexplicable error», «inválido», «profundamente equivocado»! La comunidad de la inteligencia artificial se encontraba, previsiblemente, unida en su rechazo del argumento de Penrose, aunque, desde el punto de vista de Penrose, los críticos no estaban de acuerdo

sobre cuál era el «fallo letal». Esto en sí ya era una indicación de hasta qué punto había errado el tiro, puesto que los críticos habían encontrado muchas vías diferentes de centrar el tiro sobre un gran malentendido, acerca de la propia naturaleza de la IA y su empleo de algoritmos.

2. *La Biblioteca de Toshiba*

> Sin embargo, aquellos a quienes podría parecer mejor el libro, serán probablemente los que no lo entiendan. Como biólogo evolucionista, he aprendido a lo largo de los años que la mayoría de la gente no desea verse como pesados robots programados para asegurar la supervivencia de sus genes. No creo que deseen verse tampoco como ordenadores digitales. Que alguien con impecables credenciales científicas diga que no se trata de nada de eso debe de ser tranquilizador.
>
> John Maynard Smith,
> «*What Can't the Computer Do?*»

Consideremos el conjunto de todas las máquinas de Turing, en otras palabras, el conjunto de todos los posibles algoritmos. O, más bien, para facilitar la tarea de la imaginación, consideremos en su lugar un inmenso aunque finito subgrupo de estas máquinas, las relacionadas con un lenguaje particular, y compuestas de «volúmenes» de una particular longitud: el conjunto de todas las posibles tiras de *ceros* (0) y de *unos* (1), o tiras de *bits* hasta la longitud de un megabyte (ocho millones de *ceros* y de *unos*). Consideremos que el lector de estas tiras es mi viejo ordenador portátil, un Toshiba T-1200, con sus veinte megabytes en el disco duro (se prohíbe utilizar memoria adicional, precisamente para asegurar la finitud). No debe sorprender que la inmensa mayoría de estas tiras de *bits* no hagan nada que merezca la pena mencionar, si se hace un intento de hacerlas funcionar como programas en el Toshiba. Los programas, después de todo, no son tiras aleatorizadas de *bits,* sino secuencias de *bits* cuidadosamente diseñadas, productos de miles de horas de trabajo en I+D. El programa más imaginativo que pudiera existir es algo que puede ser expresado como una tira de *ceros* (0) y de *unos* (1), y aunque mi viejo Toshiba es demasiado pequeño para hacer funcionar alguno de los verdaderamente enormes programas que han sido diseñados, es completamente capaz de hacer funcionar un amplio y representativo subgrupo de ellos: procesamiento de texto, hojas con datos numéricos en columnas y en filas, juegos de ajedrez, simulación de vida artificial, buscadores de pruebas lógicas e, incluso, unos cuantos demostradores automáticos de verdades aritméticas. Denominemos a uno de estos programas, reales o imaginados, un

programa *interesante* (en la Biblioteca de Babel esto es aproximadamente análogo a un libro legible, real o imaginario, o a un genotipo viable en la Biblioteca de Mendel). No tenemos que preocuparnos por los límites que separan lo interesante de lo no interesante; cuando tengamos dudas, lo eliminamos. No importa cómo llevemos el asunto, hay muchísimos programas muy interesantes en la Biblioteca de Toshiba aunque son tremendamente difíciles de «encontrar»; por eso las compañías de *software* hacen millonarios a unos cuantos con su *software*.

De hecho, *toda* tira de *bits* de un megabyte de longitud es un algoritmo en un sentido, precisamente en el sentido que ahora nos interesa: se trata de una receta, estúpida o inteligente, que puede seguirse mediante un mecanismo, mi Toshiba. Si tratamos tiras de *bits* de modo aleatorio, la mayor parte del tiempo el Toshiba se quedará allí emitiendo un débil zumbido (ni siquiera parpadea una luz ámbar); parafraseando a Dawkins, hay muchísimos más caminos para que llegue a ser un programa muerto que un programa vivo. Sólo un evanescente subgrupo de estos algoritmos resulta interesante de algún modo, y sólo un evanescente subgrupo de *ellos* tiene algo que ver con las verdades de la aritmética, y sólo un muy pequeño subgrupo de *estos* programas intentará generar pruebas formales de verdades aritméticas, y sólo un muy pequeño subgrupo de éstas son coherentes. Gödel nos ha demostrado que ni uno solo de los algoritmos en este subgrupo (y hay aún enormes cantidades de ellos, incluso en mi pequeño Toshiba) pueden generar pruebas de *todas* las verdades de la aritmética.

Pero el teorema de Gödel no nos dice nada acerca de otro tipo de algoritmo en la Biblioteca de Toshiba. No nos dice si hay algoritmos que puedan jugar un ajedrez aceptable. Hay, de hecho, enormes cantidades y unos pocos de estos algoritmos reales residen en mi Toshiba, y ¡nunca he conseguido ganarle a ninguno de ellos! El teorema de Gödel no nos dice si hay algoritmos que sean muy buenos para jugar a la prueba de Turing o juego de imitación. De hecho, hay uno real en mi Toshiba, una versión aligerada del famoso programa de Joseph Weizenbaum llamado ELIZA; he comprobado cómo este programa vuelve loca a gente no iniciada, que llega a la conclusión, como Edgar Allan Poe, de que allí debía estar escondido un ser humano preparando las respuestas. En un principio, me desconcertó que un ser humano cuerdo pudiera pensar que había un tipo minúsculo dentro de mi Toshiba portátil, sentado, libre, sobre una tarjeta, pero es que yo había olvidado hasta qué punto una mente persuadida puede estar llena de recursos;

allí, en mi Toshiba, se encontraría, concluían estos taimados escépticos, un *teléfono celular*.

El teorema de Gödel en particular no tiene nada que decirnos acerca de si puede haber algoritmos en la Biblioteca de Toshiba que hicieran el impresionante trabajo de «presentar como verdaderas» o «detectar como verdaderas o falsas», proposiciones posibles de aritmética. Si los matemáticos humanos pueden hacer el impresionante trabajo de «simplemente ver» con «intuición matemática» que ciertas proposiciones son verdad, quizás un ordenador pueda imitar este talento, del mismo modo que puede imitar jugar al ajedrez y mantener conversaciones: de manera imperfecta pero que no deja de ser impresionante. Esto es exactamente lo que cree la gente que trabaja en IA: que hay algoritmos heurísticos arriesgados para la inteligencia humana en general, así como los hay para jugar bien a las damas y al ajedrez, o para realizar miles de otras tareas. Y aquí es donde Penrose comete su gran error: deja de lado este conjunto de posibles algoritmos, el único grupo de algoritmos con los que tiene que ver la IA, y se concentra en un grupo de algoritmos sobre los que el teorema de Gödel nos dice realmente algo.

Los matemáticos, como dice Penrose, usan la «intuición matemática» para ver que una cierta proposición se sigue de los fundamentos de un cierto sistema. Penrose va entonces más lejos, para argumentar que no puede haber algoritmo «para» la intuición matemática o, en cualquier caso, un algoritmo práctico. Pero, entrando en este problema, Penrose pasa por alto la posibilidad de que algún algoritmo —de hecho, muchos algoritmos diferentes— puedan generar intuiciones matemáticas, aunque esta función no fuera justamente «para» lo que estaban destinados. Podemos ver claramente el error en un argumento paralelo.

El ajedrez es un juego finito (dado que hay reglas para terminar en tablas los juegos que no pueden progresar). Esto significa que hay, en principio, un algoritmo para determinar si es jaque mate o tablas; no tengo ni idea de cuál es. De hecho, puedo especificar el algoritmo de manera muy simple: 1) Dibujar un completo árbol de decisión de todos las posibles partidas de ajedrez (un número enorme pero finito). 2) Buscar el nudo final de cada juego; ganan las blancas o ganan las negras, o son tablas. 3) «Colorear» el nudo de negro, blanco o gris, según el resultado. 4) Proceda a la inversa dando un paso *completo* (un movimiento blanco más un movimiento negro) cada vez; si en el movimiento previo *todos* los caminos desde *cualquiera* de los movimientos blancos conduce a través de todas las respuestas negras a un nudo de color blanco, se colorea este nudo blanco y se mueve de nuevo hacia

atrás y así sucesivamente. 5) Hacer lo mismo para cada camino que garantice la victoria de las negras. 6) Colorear de gris todos los otros nudos. Al final de este procedimiento (cuando sea la hora de acostarse del universo) se habrán coloreado todos los nudos del árbol de todas las posibles partidas de ajedrez, retrocediendo hasta un movimiento de apertura de las piezas blancas. Ahora es el momento de jugar. Si cualquiera de los veinte movimientos está coloreado en blanco, ¡hágalo! Conduce a un garantizado jaque mate, al que se puede llegar precisamente siguiendo siempre los nudos blancos. Descarte, naturalmente, cualquier movimiento negro, dado que abre la puerta a una victoria garantizada de vuestro oponente. Si no hay movimientos blancos a la salida, escoja un movimiento gris y espere hasta que se le presente la oportunidad en la partida de un movimiento blanco. Lo peor que puede hacer son tablas. (Si todos los movimientos de apertura para las blancas están coloreados de negro, lo que es muy improbable, su única esperanza es escoger aleatoriamente y esperar que su oponente, jugando, meta la pata en un momento posterior de la partida y le deje escapar consiguiendo un camino gris o blanco).

Esto es, sin duda, un algoritmo. Ningún paso de la receta requiere intuición, y lo he especificado de forma finita y poco ambigua. El problema es que esto no es remotamente factible ni práctico, ya que el árbol del algoritmo exhaustivamente investigado es inmenso. Supongo que es interesante saber que, en principio, existe un algoritmo para jugar perfectamente al ajedrez, aunque no tenga utilidad. *Podría* existir un algoritmo *factible* para jugar un perfecto ajedrez. Gracias a Dios, nadie ha encontrado ninguna de estas exigentes características, dado que convertiría al ajedrez en un juego sin mayor interés que el juego del tres en raya. Nadie sabe si existe este algoritmo factible, pero el consenso general es que es muy improbable. No teniendo por segura su existencia, elijamos la suposición que deja en peor lugar a la IA. Supongamos que *no hay* algoritmo factible para el mate o para unas tablas garantizadas. Absolutamente ninguno.

¿Significa esto que ningún algoritmo puede conseguir un jaque mate en mi Toshiba? De ninguna manera. Como ya he confesado, los algoritmos de ajedrez de mi Toshiba no conocen la derrota contra un ser humano, es decir, contra mí. No soy muy bueno, pero espero tener tanta intuición como cualquier ser humano. Algún día podré ganarle a mi máquina si practico mucho y trabajo muy duro, aunque los programas de mi Toshiba son triviales, comparados con los actuales programas campeones. Sobre esto *puede apostar su vida* con seguridad a favor de que, en cada partida, me darán jaque mate

(aunque no a Bobby Fischer). No recomiendo a nadie que apueste su vida por la relativa excelencia de estos algoritmos —yo puedo mejorar, por lo que no desearía cargar con su muerte sobre mi conciencia— aunque, de hecho, si el darwinismo es correcto, nosotros y nuestros ancestros poseemos una serie ininterrumpida de partidas exitosas frente a otra similar de partidas fatalmente comprometidas, en los algoritmos incorporados en nuestra «maquinaria». Esto es lo que los organismos han hecho día a día desde que la vida comenzó: han apostado sus vidas a que los algoritmos que las han construido, y han actuado dentro de ellos, si se encontraban entre los afortunados organismos con cerebros, seguirán manteniéndolos vivos el tiempo suficiente para tener descendencia. La madre naturaleza nunca ha aspirado a la certeza absoluta: ella se conforma con un buen riesgo. De este modo, podemos esperar que si los cerebros de los matemáticos están haciendo funcionar algoritmos, sean algoritmos que hagan un buen trabajo en el departamento encargado de detectar las verdades, sin ser infalibles.

Los algoritmos de ajedrez de mi Toshiba, como todos los algoritmos, generan resultados garantizados, pero no están garantizados para darme jaque mate, sino, exactamente, para jugar ajedrez legal. Para eso es «para» lo que son. Del inmenso número de algoritmos garantizados para jugar ajedrez legal, unos son mejores que otros, aunque ninguno tiene la garantía de ganar contra otro; al menos esta no es la clase de cosa que uno podría probar matemáticamente, aunque, de acuerdo con el crudo hecho matemático, el estado inicial del programa x y del programa y era tal que el programa *x* ganaría todas las posibles partidas contra y. Esto significa que el siguiente argumento es falaz:

x es excelente para conseguir jaque mate;
no hay algoritmo (práctico) para dar jaque mate en ajedrez;
por lo tanto: la explicación del talento de *x* no puede ser que *x* está desarrollando un algoritmo.

La conclusión es obviamente falsa; el nivel algorítmico de la explicación es exactamente el nivel correcto desde el cual explicar el poder de mi Toshiba para ganarme al ajedrez. No es una cuestión debida a una corriente eléctrica particularmente potente conectada al Toshiba, o un secreto depósito de *élan vital* dentro de su caja plástica. Lo que le hace mejor que otros ordenadores jugando al ajedrez (yo puedo batir los más simples) es que tiene un algoritmo mejor.

Entonces, ¿qué clase de algoritmos pueden estar haciendo funcionar los matemáticos? Los algoritmos «para» *tratar de seguir vivos*. Como vimos en

el último capítulo, en nuestra consideración sobre las máquinas-robots para supervivencia, tales algoritmos tendrían una capacidad de discriminación y de planificación llena de recursos y de modo indefinido; deben ser buenos para reconocer alimentos y refugio, distinguir al amigo del enemigo, aprender a discriminar los presagios de la primavera como presagios de primavera, diferenciar los buenos de los malos argumentos e, incluso —como una especie de talento extra—, reconocer las verdades matemáticas *como* verdades matemáticas. Estos algoritmos darwinianos (Cosmides y Tooby 1989) no habrían sido diseñados naturalmente para este propósito especial, del mismo modo que nuestros ojos no fueron diseñados para diferenciar las letras *itálicas* de las negritas, aunque esto no significa que no sean tremendamente sensibles a tales diferencias, si se les da la oportunidad.

Ahora bien, ¿cómo pudo Penrose pasar por alto esta posibilidad que es retrospectivamente obvia? Es un matemático, y los matemáticos están primariamente interesados en ese evanescente subgrupo de algoritmos que pueden probar matemáticamente, que tienen poderes matemáticamente interesantes. Yo llamo a esto la visión de los algoritmos desde el ojo de Dios. Es análoga a la visión desde el ojo de Dios de los volúmenes de la Biblioteca de Babel. Podemos probar (para lo que valga la pena) que hay un volumen singular en la Biblioteca de Babel que tiene un listado, en un orden alfabético perfecto, de todos los abonados de teléfonos de la ciudad de Nueva York, cuya fortuna neta era, el 10 de enero de 1994, de más de un millón de dólares. No podía haber *tantos* millonarios propietarios de teléfonos en Nueva York, y de este modo, alguno de los posibles volúmenes de la Biblioteca debería incluir la lista de todos ellos, Pero encontrarla —o hacerla— sería una tarea empírica enorme llena de incertidumbres aun considerando que se trata de un subgrupo de nombres ya impresos en la guía real de teléfonos hasta la fecha (dejando de lado todos aquellos con números de teléfono que no están listados). Aun cuando no podemos manejar este volumen podemos nombrarlo como hicimos con la Eva Mitocondrial. Lo llamaremos *Megaphone*. Ahora podemos probar cosas acerca del *Megaphone*: por ejemplo, la primera letra impresa en la primera página es la «A», pero la primera letra de la última página no es la «A». (Este no es el estándar de las pruebas matemáticas, naturalmente, pero ¿cuáles son las posibilidades de que *ninguna* de las personas con teléfonos cuyos nombres comiencen con «A» sea un millonario?, o ¿hay solamente una página de tales millonarios en todo Nueva York?).

Como he señalado en la página 74, cuando los matemáticos piensan acerca de algoritmos lo hacen generalmente desde la perspectiva del ojo de Dios. Están interesados en probar, por ejemplo, que hay algún algoritmo con alguna interesante propiedad o que no hay tal algoritmo, y para probar estas cosas no necesitan realmente localizar el algoritmo del que hablan, extrayéndolo de un montón de algoritmos almacenados en discos blandos, por ejemplo. Nuestra inhabilidad para localizar los restos de la Eva Mitocondrial no evita que deduzcamos hechos acerca de ella. La cuestión empírica de la identificación no surge a menudo de tales deducciones formales. El teorema de Gödel nos dice que ni un solo algoritmo de los que pueden funcionar en mi Toshiba (o en otro ordenador) posee cierta propiedad matemáticamente interesante: *ser un coherente generador de pruebas de hechos aritméticos que serían generadas si se les diera bastante tiempo para funcionar*.

Esto es interesante, pero no nos ayuda mucho. Muchas cosas interesantes pueden ser probadas matemáticamente acerca de todos y cada uno de los miembros de varios grupos de algoritmos. Aplicar este conocimiento en el mundo real es otra cuestión y este es el punto negro que conduce a Penrose a pasar por alto la IA, en lugar de refutarla, como esperaba hacer. Esto se ha hecho evidente en sus posteriores intentos por reformular su tesis como respuesta a sus críticos:

Dado un algoritmo particular, este algoritmo no puede ser el procedimiento a través del cual los matemáticos humanos se cercioran de una verdad matemática. En consecuencia, los humanos no utilizan algoritmos para cerciorarse de la verdad (Penrose 1990:696).

Los matemáticos humanos no utilizan un sólido algoritmo reconocible para cerciorarse de la verdad matemática (Penrose 1991).

En la más reciente de estas respuestas a sus críticos, Penrose considera y clausura varias vías de escape de las cuales nos conciernen especialmente dos: los matemáticos pueden estar utilizando un «*irreconocible* algoritmo X horrendamente complicado» o bien un *erróneo* algoritmo Y (aunque es presumible que con algún fundamento). Penrose presenta estas vías de escape como si fueran respuestas *ad hoc* a los retos del teorema de Gödel, en lugar de hipótesis de trabajo habituales en la IA. De lo primero dice:

Esto parece estar totalmente en desacuerdo con lo que los matemáticos parecen *realmente* hacer cuando expresan sus argumentos en términos que pueden (al menos en principio) ser desmenuzados en afirmaciones que son «obvias» y en las que todos están de acuerdo. Yo consideraría esto como algo inverosímil hasta el extremo de creer que es *realmente* el horrendo e irreconocible algoritmo X, más que esos simples y obvios *ingredientes* (la cursiva es añadida), lo que se encuentra detrás de toda nuestra comprensión matemática (Penrose 1991).

Estos «ingredientes», naturalmente, los manejamos de un modo *aparentemente* no algorítmico, aunque este hecho fenomenológico es engañoso. Penrose presta una cuidadosa atención a lo que parece ser un matemático, pero pasa por alto una posibilidad —naturalmente, una probabilidad— que es familiar a los investigadores de la inteligencia artificial: la posibilidad de que *implícita* en nuestra capacidad general para tratar con tales «ingredientes» haya un programa heurístico de una complejidad que nos deja mentalmente perplejos. Este complicado algoritmo nos *aproximaría* a la competencia del perfecto entendedor y sería «invisible» para su beneficiario. Cada vez que decimos que resolvemos algún problema por «intuición» lo que realmente queremos decir es que *no sabemos cómo* lo hemos resuelto. El método más simple para hacer un modelo de «intuición» en un ordenador es sencillamente negar al programa de ordenador cualquier acceso a su propio trabajo interno. Cuando resuelve un problema y le preguntamos cómo lo ha resuelto, debe responder: «No lo sé; la solución me llegó por intuición» (Dennett 1968).

Penrose descarta su segunda vía de escape (el algoritmo erróneo): «Los matemáticos requieren un grado de rigor que hace inaceptables a estos agentes heurísticos, así que los procedimientos conocidos de este tipo no pueden ser los métodos con los que los matemáticos realmente operen» (1991).

Éste es un error más interesante, con el cual Penrose plantea la perspectiva de que la crucial prueba empírica no sería situar a un *único* matemático «en la caja» sino ¡a toda la comunidad matemática! Penrose se da cuenta de la importancia teórica del poder añadido que los matemáticos humanos consiguen acumulando sus recursos, comunicándose entre ellos hasta llegar a ser una especie de mente gigante singular que es enormemente más fiable que cualquiera de los homúnculos que podríamos tener encerrados en la caja. No es que los matemáticos tengan *cerebros* más fantásticos que el resto de los humanos (o que los chimpancés), sino que tienen herramientas mentales: las instituciones sociales en las que los matemáticos presentan sus pruebas controlan los resultados de otros colegas, cometen errores en público y cuentan con este público para corregir sus errores. Todo esto otorga a la comunidad de matemáticos poderes para discernir la verdad matemática que empequeñecen los poderes de cualquier cerebro individual humano (incluso un cerebro individual con unos periféricos papel y lápiz, una calculadora manual o un ordenador portátil), pero esto no demuestra que las mentes humanas no sean dispositivos algorítmicos; por el contrario, muestra cómo las

grúas de la cultura pueden hacer uso de los cerebros humanos en la distribución de procesos algorítmicos, sin límites discernibles.

Penrose no lo ve así. Llega a decir que «es nuestra habilidad general para *comprender* (una habilidad no algorítmica)» la que explica nuestras habilidades matemáticas, y entonces concluye: «No fue un *algoritmo x* el que fue favorecido en el hombre (al menos) por la selección natural, sino esta admirable capacidad para comprender» (Penrose 1991). Aquí él comete la falacia que precisamente expuse mediante el ejemplo del ajedrez. Penrose arguye:

x puede comprender;
no hay algoritmo factible para comprender;
por consiguiente: lo que la selección natural ha seleccionado, cualquier cosa que sea lo que nos hace entender, no es un algoritmo.

Esta conclusión es un *non sequitur*. Si la mente es un algoritmo (contrariamente a lo que sostiene Penrose), seguramente no es un algoritmo que sea reconocible o accesible a aquéllos cuyas mentes creó. Es, en sus propios términos, irreconocible. Como un producto de procesos de diseños biológicos (genéricos e individuales) es, casi con certeza, uno de aquellos algoritmos que se encuentran en uno u otro lugar del enorme espacio de los algoritmos interesantes, lleno de errores tipográficos o «fallos», pero lo bastante bueno para apostar nuestra vida por él, hasta ahora. Penrose considera esta posición como una posibilidad traída por los pelos, pero si esto es todo lo que él puede decir en contra, lo que sucede es que todavía no se ha enfrentado seriamente con la mejor versión de la inteligencia artificial «fuerte».

3. *El imaginario ordenador de la gravedad cuántica: lecciones desde Lapland*

> Creo firmemente en el poder de la selección natural, pero no veo cómo la selección natural, en sí misma, puede desarrollar algoritmos que pudieran poseer el tipo de juicios conscientes de la validez de otros algoritmos que parece que tenemos.
>
> Roger Penrose,
> *The Emperor's New Mind Concerning Computers, Minds, and the Laws of Physics*

> No creo que el cerebro haya llegado a ser lo que es según explicó Darwin. De hecho, es improbable. Simples mecanismos no pueden generar el cerebro. Creo que los

elementos básicos del universo son simples. La fuerza de la vida es un elemento primitivo del universo y obedece a ciertas leyes de acción. Estas leyes ni son simples ni mecánicas.

Kurt Gödel[157]

Cuando Penrose insiste en que el cerebro no es una máquina de Turing, lo que importa es comprender lo que no dice. No hace la obvia afirmación (y obviamente irrelevante) de que la original máquina de pensar de Turing no es un modelo de cerebro: un pequeño artilugio colocado a horcajadas en una tira de papel, que examina cada vez un cuadrado de esta tira. Nadie ha pensado otra cosa. Tampoco dice Penrose que el cerebro no es un ordenador en serie —una «máquina de Von Neumann»— sino, más bien, un masivo ordenador en paralelo. Penrose no dice tampoco que el cerebro utilice la aleatoriedad o la pseudoaleatoriedad al hacer funcionar sus algoritmos. Penrose cree —aunque algunos no piensan así— que los algoritmos que sacan partido de amplias dosis de aleatoriedad siguen siendo algoritmos, dentro del territorio que incumbe a la inteligencia artificial, y siguen cayendo bajo las limitaciones que el teorema de Gödel aplica a todas las máquinas de Turing, de cualquier tamaño o forma[158].

Además, y como secuela de los comentarios que su libro provocó, Penrose da ahora por probado que los programas heurísticos son también algoritmos y, reconocido esto, si persiste en su intento de hallar un argumento contra la IA, ha de conceder su tremendo poder para seguir las huellas de las verdades de la aritmética y de cualquier otro tipo, si no con perfección, al menos de modo impresionante. Penrose ofrece una clarificación adicional: cualquier ordenador que opera mediante interacciones con un medio externo es un ordenador algorítmico, *en el supuesto de que el mismo medio ambiente sea enteramente algorítmico*. (Si los ganchos celestes crecieran como setas —o, para acentuar más nuestro argumento, como oráculos colocados sobre las setas— y un ordenador recibiese ayuda a través de una ocasional comunicación con estos ganchos celestes, entonces lo que hiciera ese ordenador no sería un algoritmo).

Ahora, una vez que se han presentado las clarificaciones pertinentes, ¿qué es lo que sostiene Penrose? En mayo de 1993, pasé una semana con Penrose y con algunos físicos suecos y otros científicos, en una discusión de nuestros diferentes puntos de vista sobre estas materias, organizada como reunión de trabajo, celebrada en Abisko, una estación de investigación de la tundra, situada muy al norte de Suecia, en el círculo polar ártico. Quizás el sol de

media noche podría ayudarnos, tanto como nuestros huéspedes suecos, a iluminar el camino, pero, en cualquier caso, creo que tanto Penrose como yo quedamos lejos de ser iluminados. Penrose propuso una revolución en la física, centrada en una nueva —y aún no formulada— teoría de la «gravedad cuántica», que explicaría el modo según el cual el cerebro humano trasciende las limitaciones de los algoritmos. ¿Está pensando Penrose en un cerebro humano con los poderes especiales de la física cuántica, sean éstos un gancho celeste o una grúa? Esta era la pregunta con la que yo marché a Laponia, Suecia, en busca de una respuesta, y la respuesta con la que regresé fue ésta: definitivamente, Penrose había estado buscando un gancho celeste. Yo creo que Penrose habría aceptado una nueva grúa, aunque dudo que encontrara alguna.

Descartes y Locke, y más recientemente Edgar Allan Poe, Kurt Gödel y J. R. Lucas, pensaron que la alternativa a una mente «mecánica» sería una mente *inmaterial*, o un alma, según la tradición. Hubert Dreyfus y John Searle, los más recientes escépticos con respecto a la IA, han descartado tal dualismo y opinan que la mente es naturalmente el cerebro, pero que el cerebro no es un ordenador *ordinario*; tiene «poderes causales» (Searle 1985) que van más allá de hacer funcionar algunos algoritmos. Ni Dreyfus ni Searle han sido muy explícitos respecto a lo que puedan ser estos poderes especiales, o cuál de las ciencias físicas podía ser la correcta para explicarlos, pero otros se han preguntado si la física podría tener la clave. Para muchos de ellos, Penrose aparece como un caballero con una brillante armadura.

¡La física cuántica al rescate! A lo largo de los años, se han avanzado varias propuestas diferentes acerca de cómo el poder de los efectos cuánticos puede ser utilizado para dotar al cerebro de poderes especiales, más allá de los que posee un ordenador ordinario. J. R. Lucas se sintió muy atraído por la introducción de la física cuántica en este campo, aunque pensó que los hiatos de indeterminación de la física cuántica podrían permitir a un espíritu cartesiano intervenir, agitando las neuronas, con tal de conseguir algún poder mental extra fuera del cerebro, una doctrina que ha sido enérgicamente defendida por sir Jonh Eccles, el neurofisiólogo laureado con el Premio Nobel, quien durante años escandalizó a sus colegas con su desenfadado dualismo (Eccles 1953, Popper y Eccles 1977). No es este el momento, ni el lugar, de exponer las razones para rechazar este dualismo —el momento y el lugar se encuentran en Dennett [1991a, 1993d]—, dado que Penrose rechaza el dualismo con tanto vigor como cualquiera que milite en el campo materialista. De hecho, lo que es alentador en los ataques de Penrose a la IA

es la insistencia en su esperanza de reemplazarla con algo que podría ser una ciencia física de la mente, no algún misterio inexplorable que tuviera lugar en la «tierra de irás y no volverás» del dualismo.

Sin abandonar la esfera física, sería posible conseguir algunos extraños nuevos poderes que tuvieran su origen en las partículas subatómicas, según una reciente especulación sobre «ordenadores cuánticos» (Deutsch 1985). Se afirma que este ordenador cuántico se aprovecharía de la «superposición de *eigenstates*» antes del «colapso del paquete de ondas» para examinar cuidadosamente inmensos (sí, inmensos) espacios de búsqueda, en cantidades ordinarias de tiempo. Por tratarse de una especie de ordenador paralelo supermasivo, podría hacer inconmensurablemente muchas cosas «al mismo tiempo», lo que haría factibles todas las clases de algoritmos que, de otra manera, no lo serían, como el algoritmo para jugar el ajedrez perfecto. Sin embargo, esto *no* es lo que Penrose está buscando para tales ordenadores, porque incluso si fueran posibles, no dejarían de ser máquinas de Turing y, en consecuencia, tendrían capacidad únicamente para computar las funciones oficialmente computables: los algoritmos (Penrose 1989:402). Por lo tanto, estos ordenadores quedarían bajo las limitaciones descubiertas por Gödel. Penrose está buscando un fenómeno que sea verdaderamente *no computable*, no sólo poco práctico de computar.

Reconoce Penrose que la física actual (con inclusión de la actual física cuántica) es *toda* ella computable, pero piensa que hemos de revolucionar la física, incorporando una teoría de la «gravedad cuántica», explícitamente no computable. ¿Por qué piensa Penrose que esta teoría (que ni él ni nadie ha formulado todavía) tendría que ser no computable? Porque, de otro modo, la inteligencia artificial es posible, y él cree que ya ha demostrado, a través de la argumentación del teorema de Gödel, que la IA no es posible. Esto es todo. Penrose admite cándidamente que ninguna de sus razones para creer en la no computabilidad de la teoría de la gravedad cuántica salen de la propia física cuántica; la única razón que tiene para pensar que una teoría de la gravedad cuántica sería no computable es, después de todo, que de otra manera la IA sería posible. En otras palabras, Penrose tiene la corazonada de que algún día encontraremos un gancho celeste. Esta es la corazonada de un brillante científico, aunque él mismo admite que se trata solamente de una corazonada.

En una revisión del reciente libro del físico Steven Weinberg *El sueño de una teoría final* (recuerden que en el capítulo 3 Weinberg lanzó dos hurras al reduccionismo), Penrose medita así:

> Desde mi punto de vista, si ha de haber una teoría final, ésta sólo puede ser un esquema de una naturaleza muy diferente. Más que una teoría física en su sentido ordinario, sería algo así como un principio, un principio matemático cuya implementación podría en sí misma implicar sutilezas no mecánicas (y quizás, incluso, creatividad) (Penrose 1993:82).

Así que no es sorprendente que Penrose se haya expresado con profundo escepticismo sobre el darwinismo. Y los fundamentos de su posición nos son ya familiares: no puede imaginar cómo la «selección natural de algoritmos» podría hacer todo este buen trabajo:

> Hay serias dificultades con la descripción de cómo los algoritmos se mejoran supuestamente a sí mismos de este modo. Esto ciertamente no funcionaría siguiendo las especificaciones de una máquina de Turing normal, dado que parece cierto que una «mutación» convertiría a la máquina en totalmente inútil, en lugar de modificarla tan sólo ligeramente (Penrose 1990:654).

La mayoría de las mutaciones, piensa Penrose, son invisibles para la selección natural o fatales; tan sólo unas pocas mutaciones mejoran las cosas. Esto es cierto, pero es exactamente tan cierto para los procesos evolutivos que producen las mandíbulas de los cangrejos, como para los procesos que producen los estados mentales de los matemáticos. La convicción de Penrose de la existencia de estas «serias dificultades» queda devaluada, como le sucedió a la convicción de Poe, por el crudo hecho histórico de que los algoritmos genéticos y de su clase están diariamente superando estas aterradoras probabilidades y mejorándose a sí mismos, mediante saltos y limitaciones (en la escala geológica del tiempo).

Si nuestros cerebros *estuvieran* equipados con algoritmos, argumenta Penrose, la selección natural tendría que haber diseñado aquellos algoritmos, pero:

> Las «sólidas» especificaciones son las *ideas* que sirven de fundamento a los algoritmos. Pero las ideas son cosas que, donde sabemos, necesitan mentes conscientes para sus manifestaciones (Penrose 1989:415).

En otras palabras, el proceso de diseño tendría que apreciar, de algún modo, la racionalidad de aquellos algoritmos que estaba diseñando y ¿no corresponde esto a una mente consciente? ¿Puede haber razones reconocidas sin el reconocimiento de alguna mente consciente? Sí, dijo Darwin, puede haberlas. La selección natural es el relojero *ciego*, el relojero *inconsciente*, pero, al mismo tiempo, un descubridor de movimientos forzados y otras buenas mañas. Esto no es tan inconcebible como muchos lo han considerado.

> En mi opinión, hay aún algo misterioso acerca de la evolución con su aparente «incierto caminar» hacia algún futuro propósito. Las cosas, al menos, *parecen* organizarse a sí mismas, de alguna manera, mejor

de lo que «debieran» al tratarse de una evolución ciega y al azar, y de una selección natural. Puede ser que tales apariencias sean completamente falaces. Podría tratarse, más bien, de algo parecido a un funcionamiento regido por las leyes de la física, lo que permitiría a la selección natural ser un proceso mucho más efectivo de lo que sería con leyes arbitrarias (Penrose 1989:416).

No podría encontrarse una manifestación de Penrose más clara y sincera que la arriba citada, de su esperanza en encontrar ganchos celestes. Y aunque no podemos descartar «en principio» la existencia de un gancho celeste representado por la teoría de la gravedad cuántica, Penrose no ha proporcionado todavía una razón para creer en ella. Si esta teoría de la gravedad cuántica fuera ya una realidad, sería realmente una grúa, pero no ha sido así hasta ahora, y dudo de que lo sea. Al menos lo está intentando, Penrose desea que esta teoría aporte una visión científica unificada de cómo trabaja la mente, no una excusa para declarar a la mente una impenetrable fuente última de significados. Mi propia opinión es que el camino que Penrose está explorando ahora —en concreto, los posibles efectos cuánticos que ocurrirían en los microtúbulos del citoesqueleto de las neuronas, una idea entusiásticamente promovida por Stuart Hameroff en Abisko— es un caballo retirado de la carrera y, además, no es un tema apropiado para esta ocasión. (No puedo resistir la tentación de plantear a Penrose una pregunta, para que la considere cuidadosamente: si la magnífica propiedad cuántica se encuentra oculta en los microtúbulos de las neuronas, ¿no significa esto que las cucarachas poseen también mentes no computables? Tienen la misma clase de microtúbulos que nosotros).

Si un cerebro con gravedad cuántica, al estilo de Penrose, fuese verdaderamente capaz de actividad no algorítmica, y si existieran tales cerebros, y si nuestros cerebros son ellos mismos el producto de un proceso algorítmico evolutivo, surge una curiosa inconsistencia: un proceso algorítmico (selección natural en sus varios niveles y encarnaciones) crea un subproceso no algorítmico o subrutina lo que convierte el proceso *global* (evolución hasta *incluir* los cerebros humanos de los matemáticos) en un proceso no algorítmico, después de todo. ¡Sería como una cascada de grúas creando, eventualmente, un gancho real en el cielo! No es extraño que Penrose tenga sus dudas acerca de la naturaleza algorítmica de la selección natural. Si fuera, verdaderamente, un proceso algorítmico a todos los niveles, todos sus productos deberían ser también algorítmicos. En mi opinión, esta no es una contradicción formal ineludible; Penrose puede encogerse de hombros y proponer que el universo contiene estos núcleos básicos de poder no algorítmico, no creados a sí mismos por selección natural en ninguna de sus variantes, aunque puedan ser incorporados por los dispositivos algorítmicos

como objetos encontrados cuando quiera que sean encontrados (como la batracomancia). Aquellos serían verdaderamente ganchos celestes no degradados.

Creo que esta posición es posible, aunque Penrose debe enfrentarse con una embarazosa escasez de evidencias a su favor. El físico Hans Hansson llegó a Abisko con un buen reto, comparando una máquina de movimiento continuo con un ordenador detector de verdades. Diferentes ciencias, señaló Hansson, pueden ofrecer diferentes atajos a veredictos acerca de proyectos. Si alguien se presentara al gobierno sueco con un plan para construir una máquina de movimiento perpetuo (con financiación del gobierno), Hansson testificaría sin dudarlo, como físico, que este proyecto sería —*tendría que ser* — tirar el dinero del gobierno. No podría tener éxito porque la física ha demostrado que el movimiento perpetuo es imposible. ¿Acaso pensó Penrose que él había ofrecido un *tipo* de prueba similar? Si algún empresario en el campo de la inteligencia artificial acudiera al gobierno para pedir dinero para construir una máquina detectora de verdades matemáticas, ¿estaría Penrose dispuesto a declarar que eso sería tirar el dinero?

Para hacer más específica la cuestión, consideremos algunas variedades especiales de verdad matemática. Es bien conocido que no puede existir un programa de todo uso que pueda examinar cualquier otro programa y decir si existe o no en éste un bucle infinito por lo que, en consecuencia, no se detendrá si se inicia. Esto es conocido como el *problema de la detención* y existe una prueba, estilo Gödel, de que es insoluble. (Este es uno de los teoremas a los que Turing aludía en su comentario de 1946, citado al inicio de este capítulo). Ningún programa, que en sí mismo está garantizado de terminar, puede decir de todo programa (finito), si terminará o no terminará. Pero, incluso puede ser práctico —merecería la pena su coste económico— tener disponible un programa que fuera muy, muy bueno (si no perfecto) para esta tarea. Otro tipo de problemas interesantes los plantean las conocidas *ecuaciones diofantinas,* y sabemos que no hay algoritmos con la garantía de resolver estas ecuaciones. Si nuestras vidas dependiesen de estas soluciones, ¿deberíamos gastarnos unas monedas en un programa para resolver las ecuaciones diofantinas «en general» o para analizar el problema de la detención «en general»? (Recuérdese: no debemos gastarnos ni una peseta en las máquinas de movimiento perpetuo, ni siquiera para salvar nuestras vidas, dado que sería tirar el dinero en una tarea imposible).

La respuesta de Penrose fue esclarecedora: si los candidatos para el control de la verdad «de algún modo burbujean sobre el terreno» entonces

sería inteligente gastar el dinero, pero si algún agente inteligente es la fuente de los candidatos y examina el programa en nuestro controlador de la verdad, entonces puede anular a nuestro algorítmico controlador de la verdad mediante la construcción del candidato (o candidatos) «erróneo»; una ecuación insoluble por éste, o un programa cuyas perspectivas de terminación le confundirían. Para que esta distinción sea más gráfica, podemos imaginar al pirata del espacio, Rumpelstiltskin es su nombre, reteniendo el planeta como rehén, pero liberándonos indemnes si podemos responder a mil preguntas sobre si proposiciones aritméticas son verdaderas o falsas. ¿Debemos colocar a un matemático humano en el estrado de los testigos, o bien a un ordenador detector de verdades diseñado por el mejor programador? Según Penrose, si ligamos nuestro destino al ordenador y *permitimos que Rumpelstiltskin vea el programa* del ordenador, el pirata del espacio puede diseñar una proposición que busque un talón de Aquiles en nuestra máquina y la haga fracasar. (Esto sería verdad, independientemente del teorema de Gödel, si nuestro programa fuera un detector de verdades heurístico, que aceptara riesgos como los programas de ajedrez). Pero Penrose no nos ha dado razones para creer que esto no es exactamente tan verdad en cualquiera de los matemáticos humanos que pudiéramos llevar al estrado de los testigos. Ninguno de nosotros es perfecto e incluso un equipo de expertos tiene, sin duda, algunas debilidades que Rumpelstiltskin puede explotar, dando información sobre sus cerebros. Von Neumann y Morgenstern inventaron la teoría de los juegos para tratar el caso particular de problemas complicados que la vida nos presenta cuando hay agentes a nuestro alrededor que compiten con nosotros. Es una postura inteligente defender a nuestro cerebro de tales competidores, sean humanos u ordenadores. La razón de que un agente competitivo marque una diferencia en este ejemplo es que el espacio de todas las verdades matemáticas es inmenso, el espacio de las soluciones de la ecuación diofantina es un inmenso pero «evanescente» subespacio dentro de aquel, y las probabilidades de alcanzar una verdad de *modo aleatorio* que «rompa» o «derrote» a nuestra máquina son verdaderamente mínimas, mientras que una inteligente *búsqueda* a través de tal espacio, guiada por el conocimiento del estilo particular del oponente y sus limitaciones, sería como encontrar una aguja en un pajar: una contrajugada aplastante.

Rolf Wasen planteó otra interesante cuestión en Abisko. La clase de algoritmos *interesantes* incluye, sin duda, muchos que no son *humanamente* accesibles. Para expresarlo con un cierto dramatismo, hay programas en la Biblioteca de Toshiba que no funcionan en mi Toshiba pero que valoro

gracias al magnífico trabajo que hacen para mí, pero que no serían capaces de crear ni los programadores humanos ni sus artefactos (existen ya programas para escribir programas). ¿Cómo puede ser esto? Ninguno de estos admirables programas mide más de un megabyte, y existen ahora muchísimos programas reales que son más largos. De nuevo debemos recordar lo inmenso que es el espacio de tales posibles programas. Como el espacio de posibles novelas de quinientas páginas, de sinfonías de cincuenta minutos, o poemas de cinco mil versos, el espacio de los programas de un megabyte de largo será solamente ocupado por los más delgados hilos de la realidad, no importa cuánto nos esforcemos.

Hay novelas cortas que nadie podría escribir que no serían únicamente *best-sellers*; serían instantáneamente reconocidas como clásicos. El tecleado necesario para escribirlas está disponible en mi procesador de textos, y el número total de pulsaciones de teclas en tal libro es trivial, pero incluso está más allá del horizonte de la creatividad humana. Cada particular creador, cada novelista o compositor o programador de programas para ordenadores, recorre el espacio de diseño con un particular conjunto de hábitos conocido como *estilo* (Hofstadter 1985:111). Es el estilo el que nos limita y nos capacita, dando una dirección positiva a nuestra exploración pero solamente para presentarnos diferentes regiones vecinas fuera de nuestros límites; si están fuera de nuestros límites en particular, probablemente están fuera de límites para todos y para siempre. Los estilos individuales son verdaderamente únicos, el producto de incalculables miles de millones de encuentros inesperados y fructíferos a lo largo de las edades, encuentros que han producido primeramente un único genoma, y, después, una única cría y, finalmente, un único conjunto de experiencias vitales. Marcel Proust nunca tuvo la oportunidad de escribir una novela sobre la guerra de Vietnam, y nadie podrá nunca escribir *esa* novela relatando *esa* época a *su* manera. Estamos enclavados, por nuestra realidad y por nuestra finitud, en un insignificante rincón del espacio total de posibilidades, pero ¡qué hermosa realidad no es accesible, gracias al trabajo de I+D de todos nuestros predecesores! Podemos también incrementar lo que tenemos, con lo cual dejaremos más cosas a nuestros descendientes, para que trabajen con ellas.

Es el momento de devolver al contrincante el peso de la prueba, ese método que Darwin utilizó cuando retó a sus críticos a describir alguna *otra* vía —otra que no fuera la selección natural— en la cual podían haber surgido todas las maravillas de la naturaleza. Aquellos que piensan que la mente humana es no algorítmica deben considerar la soberbia que supone esta

convicción. Si la peligrosa idea de Darwin es correcta, un proceso algorítmico es lo suficientemente potente para diseñar un ruiseñor y un árbol. ¿Sería bastante más difícil para un proceso algorítmico escribir una oda a un ruiseñor o un poema tan hermoso como un árbol? Seguramente la segunda regla de Orgel es correcta: la evolución es más inteligente que nosotros.

CAPÍTULO 16
Sobre el origen de la moral

I. «*¿E pluribus unum?*»

> La naturaleza (el arte a través del cual Dios ha hecho el mundo y lo gobierna) es también imitada, como en otras muchas cosas, por el arte del hombre que puede construir un animal artificial. Si la vida no es sino un movimiento de extremidades, cuyos comienzos se encuentran en alguna parte interna principal, ¿por qué no podemos decir que todos los autómatas (ingenios que se mueven a sí mismos mediante muelles y ruedas como hace un reloj) tienen una vida artificial? ¿Qué es el corazón, sino un muelle? ¿Qué son los nervios, sino muchas cuerdas? ¿Qué son las articulaciones, sino muchas ruedas que dan movimiento a la totalidad del cuerpo, tal como fue el propósito del Artífice? El arte va todavía más lejos, hasta imitar ese racional y más excelente trabajo de la naturaleza: el hombre. Por el arte es creado ese gran LEVIATÁN llamado REPÚBLICA O ESTADO (en latín CIVITAS), que no es sino un hombre artificial; aunque de mayor estatura y fuerza que el natural, para cuya protección y defensa fue destinado; y en el cual la soberanía es un alma artificial, que da vida y movimiento a la totalidad del cuerpo.
>
> Thomas Hobbes, *Leviatán*

Thomas Hobbes fue el primer sociobiólogo, doscientos años antes que Darwin. Como las palabras iniciales de su obra maestra dejan claro, él veía la creación del estado como si se tratara fundamentalmente de un artefacto que fabricara a otro, una especie de vehículo para la supervivencia de un grupo, «concebido» para la «protección y defensa» de sus ocupantes. El frontispicio de la edición original muestra lo seriamente que Hobbes se tomó su propia metáfora.



Figura 16.1

Sin embargo, ¿por qué califico a Hobbes de sociobiólogo? En realidad, él no tuvo ocasión de plantearse, como los sociobiólogos de hoy en día, la utilización de las ideas de Darwin en un análisis de la sociedad. Sin embargo, Hobbes atisbo, con claridad y confianza, la fundamental tarea darwiniana: vio que *había* de ser contada una historia acerca de cómo el estado, en primer lugar, llegó a ser creado, y cómo esto trajo consigo algo completamente nuevo en la faz de la Tierra: la moral. Sería una historia que nos trasladaría desde un tiempo en el cual no estaba claramente definido lo correcto y lo erróneo, sino una competición amoral, hasta un tiempo en el cual de modo manifiesto existía lo correcto y lo erróneo (en algunas partes de la biosfera), a través de un proceso que gradualmente introdujo las características

«esenciales» de una perspectiva ética. Dado que el período relevante era prehistórico y no había huellas fósiles que consultar, la historia debería ser una reconstrucción racional, exactamente del tipo de esas que hemos denominado una especie de «historias, al fin y al cabo» como las de Kipling (con lo que cometemos otro anacronismo).

Erase una vez un tiempo, dijo Hobbes, en el que no existía moral alguna. Había vida; había seres humanos, que incluso poseían lenguaje, de modo que tenían memes (con lo que cometemos otro anacronismo). Cabe presumir que tenían palabras —y en consecuencia memes— para lo bueno y lo malo, pero no para lo *éticamente* bueno y malo. «Para las nociones de lo correcto y lo erróneo, de la justicia y de la injusticia no había lugar». Así, mientras distinguían una buena espada de una mala, una buena cena de una mala, un buen cazador (un experto asesino de lo que sería la cena) de un mal cazador (que ahuyentaba la presa), no tenían concepto de una persona buena o justa, una persona moral, o un buen acto, un acto moral, o sus contrarios, villanos y vicios. Podían apreciar que algunos eran más peligrosos que otros o mejores luchadores, o parejas más deseables, pero sus perspectivas no iban más allá de esto. No poseían el concepto de correcto y erróneo debido a que «hay cualidades que relacionan a los hombres en sociedad, pero no en soledad». Hobbes denominó a esa época en nuestra prehistoria «el estado de la naturaleza», debido a su semejanza, en sus más importantes aspectos, con la lucha de todos los demás animales en la selva, hasta ese día. En el estado de la naturaleza «no hay lugar para la industria; debido a que el fruto de eso es incierto… ni artes; ni letras, ni sociedad; y lo que es peor de todo, miedo continuo y peligro de muerte violenta; y la vida del hombre, solitaria, pobre, repulsiva, brutal y corta».

Entonces, un hermoso día, sucedió una mutación. Un día, cuando surgió otro conflicto, como venía sucediendo hasta entonces, algo nuevo sucedió. En lugar de persistir en las miópicas conductas egoístas de deserciones y desconfianzas mutuas que habían reinado hasta entonces, estos competidores especialmente afortunados lanzaron una nueva idea: cooperación para beneficio mutuo. Hicieron un «contrato social». Mientras que antes había familias, hordas o tribus, ahora había nacido un diferente *tipo* de grupo, una sociedad. Fue el nacimiento de la civilización. Y el resto, como suele decirse, es historia.

¡Cómo habría admirado Hobbes la historia de Lynn Margulis acerca de la revolución eucariota, y la creación, de esa forma, de la vida multicelular! Mientras que antes no había nada más que aburridos procariotas,

arrastrándose a lo largo de sus desagradables, toscas y cortas vidas, ahora podría haber organismos multicelulares, los cuales, gracias a la división del trabajo entre un grupo de células especializadas, podrían implicarse en la industria (en particular, un metabolismo basado en el oxígeno) y las artes (percepción y locomoción de largo alcance, y coloración protectora, etc.). Y, a su debido tiempo, sus descendientes crearían sociedades multicelulares de un tipo muy peculiar, conocidas (hasta recientemente) como hombres, capaces de crear letras (o representaciones), con las cuales ellos se daban a intercambios de forma promiscua; esto hizo posible una segunda revolución.

¡Cómo hubiera admirado Hobbes la historia de Richard Dawkins acerca del nacimiento de los memes y la creación, de ese modo, de las personas, que ya no eran meros vehículos de supervivencia para sus genes! Estas historias, compuestas mucho tiempo después de Hobbes, narraban los grandes pasos en la evolución que antecede al paso que él decide describir: el paso de las personas desprovistas de moral a los ciudadanos. Hobbes interpretó correctamente este paso, como un gran paso en la historia de la vida en este planeta, y decidió contar, como mejor pudiera, la historia de las condiciones bajo las cuales este paso pudo ser dado, y una vez cumplido, *evolutivamente reforzado* (para utilizar un anacronismo más). Aunque esto no fue un salto sino un pequeño paso, tendría enormes consecuencias, porque era naturalmente el nacimiento de un monstruo prometedor.

Sería un error leer a Hobbes como a un presunto historiador que estuviera simplemente especulando de manera irresponsable. Sabía, seguramente, que no había esperanza de encontrar el lugar del nacimiento de la civilización con las herramientas de la historia (o la arqueología, una disciplina todavía no inventada), pero esta no es la cuestión. Sin duda, la real secuencia prehistórica de los acontecimientos fue más confusa, y enrevesada, con elementos de una casi sociedad (del tipo de la que vemos entre las hordas de ungulados y la manada de predadores) un casi lenguaje (del tipo que vemos entre las llamadas de alarma de los pájaros y de los monos e incluso entre las abejas conductoras) y quizás elementos de casi moral (del tipo puesto en evidencia en los monos[159], así como en las solícitas ballenas y en los delfines). La reconstrucción racional de Hobbes fue una gran hipersimplificación, un modelo que tenía como objetivo ilustrar los principios esenciales, mientras que dejaba de lado los detalles sucios y no identificables. Y, sin ninguna duda, la exposición de Hobbes era demasiado simple, incluso en sus propios términos. Hoy, en la estela de cientos de investigaciones en los escondrijos y en las grietas de la teoría del juego, las competiciones sobre el dilema del

prisionero, y cosas parecidas, sabemos que Hobbes era demasiado optimista, demasiado sanguíneo (por utilizar una palabra de su vocabulario) sobre las condiciones bajo las cuales un contrato social sería evolutivamente reforzable. Pero él fue el explorador pionero de este fenómeno.

Siguiendo sus pasos, Jean-Jacques Rousseau y varios pensadores ingleses, con la inclusión de John Locke, ofrecieron su propia reconstrucción racional del nacimiento de la sociedad. Historias «contractuales» más complicadas han sido desarrolladas en años recientes. La más famosa y sofisticada es la *Teoría de la justicia* (1971) de John Rawls, aunque hay otras. Todas están de acuerdo en aceptar que la moral es, de un modo o de otro, un producto emergente de una mayor innovación en perspectiva que había sido conseguida justamente por una especie, el *Homo sapiens*, sacando partido de ese único medio extra de transferencia de información que es el lenguaje. En el experimento mental de Rawls acerca de cómo una sociedad debe formarse, imaginamos un tiempo, en el nacimiento de la sociedad, cuando los habitantes se reunían para considerar qué tipo de diseño tendría su sociedad. Se reunían para razonar juntos acerca de eso, hasta que consiguieron lo que Rawls llama un «equilibrio reflectivo», un acuerdo estable que no puede ser desbaratado por una consideración posterior En este sentido, la idea de Rawls es, como la idea de Maynard Smith, una estrategia evolucionada estable o ESS (Evolutionary Stable Strategy), pero con una importante diferencia: se trata de *gente* que hace el cálculo, no de pájaros o de pinos, ni de otros simples competidores en los juegos de la vida. La innovación clave en el escenario de Rawls, diseñado para asegurarse de que un indebido carácter egoísta entre los participantes en este ejercicio de reflexión, lo reduzca a la nada, es lo que él llama el «velo de ignorancia». Todo el mundo se encuentra en posición de votar en favor de un diseño de sociedad, pero cuando decidimos en qué sociedad seríamos felices viviendo y prestándole nuestra fidelidad, votamos sin conocer cuál será nuestro papel en ella o en qué nicho de su espacio vital estaremos situados. Puede que sea el de un senador, un cirujano, un barrendero o un soldado; no lo sabremos hasta que hayamos votado. La elección tras el velo de la ignorancia asegura que la gente preste la consideración debida a los efectos probables, a los costes y beneficios, para toda la ciudadanía, con inclusión de aquellos que son, con mucho, los peores.

La teoría de Rawls ha recibido, y merecido, más atención que ningún otro trabajo de ética en este siglo, y como es usual, estoy presentando una versión hipersimplificada del tema. Mi idea es atraer la atención sobre la ubicación de este trabajo, y de todo el trabajo que ha provocado e inspirado, en relación

con el pensamiento darwiniano en general, y con la ética evolucionista, en particular. Cabe señalar, especialmente, que mientras que Hobbes presentaba una reconstrucción racional de algo que realmente sucedió —algo que debía haber sucedido— Rawls presenta un experimento mental acerca de lo que, si sucediera, sería lo *correcto*. El proyecto de Rawls no es historia especulativa o prehistoria, sino un proyecto enteramente normativo: un intento de demostrar cómo *debe* responderse a las cuestiones éticas y, muy especialmente, un intento de *justificar* un conjunto de normas éticas. Hobbes esperaba resolver el problema normativo acerca de lo que la ética *debe* ser —el problema de Rawls— pero, como era un reduccionista insaciable, trató de matar dos pájaros de un tiro: deseaba explicar cómo lo correcto y lo erróneo llegaron a existir en primer lugar, un ejercicio de imaginación al modo darwiniano. No es necesario decir que la vida es mucho más complicada que todo eso, pero fue un bello intento.

El discurso de Hobbes en el *Leviatán* tiene un hermoso timbre panglossiano, en los dos readaptados de esta popular palabra. En primer lugar, en cuanto presupone la racionalidad (o prudencia, como él la llama) de los agentes cuya mutua solución se supone que es la sociedad, Hobbes considera el nacimiento de la sociedad como dictado por la razón, un movimiento forzado o, al menos, fuertemente auspiciado por la razón, una buena maña. En otras palabras, el relato de Hobbes es algo así como una historia adaptacionista, y no le ha dañado en absoluto. Pero, en segundo lugar, al apelar, como hace, a nuestro sentido del *bien de nuestra propia especie*, está en condiciones de adormecernos con modelos excesivamente optimistas de cómo esto ha llegado a suceder, y esta es una grave crítica. Pudo ocurrimos que, de cualquier modo que sucediera, el nacimiento de la moral fuera una buena cosa *para nosotros*, pero debemos procurar no ser indulgentes con ese tipo de reflexión. Independientemente de hasta qué punto fue verdad, esto no puede explicar cómo estas prácticas a las cuales estamos retrospectivamente tan agradecidos, llegaron a existir y persistieron. La racionalidad del grupo *puede no* ser asumida, no más de lo que podemos asumir que dado que nos hemos beneficiado poderosamente de la revolución eucariota, es explicado de esta manera. La racionalidad del grupo, o cooperación, ha de ser alcanzada, y esta es una gran tarea de diseño, si estamos considerando alianzas de procariotas o alianzas de nuestros más recientes ancestros. De hecho, una gran parte de los mejores trabajos sobre ética en los años recientes, se ocupa precisamente de este tema (Parfit 1984, Gauthier 1986, Gïbbard 1985).

Antes de examinar más de cerca, en este aspecto, la dificultad humana, cabe considerar con más cautela la metáfora que Hobbes nos invita a aceptar seriamente, con la ayuda de la mejorada perspectiva aportada por la revolución darwiniana que ha acontecido. ¿En qué aspectos es la sociedad algo así como un organismo gigante, y en qué aspectos es diferente?

Los organismos multicelulares han resuelto el problema de la solidaridad del grupo. No escuchamos nunca historias del pulgar de un persona alzándose en guerra civil contra los dedos vecinos, o de las alas de un águila que han ido a la huelga, rehusando trabajar a menos que se les haga alguna concesión, por parte del pico o (más a favor de mi argumento) de las gonadas. Y, ahora que tenemos la perspectiva del ojo del gen desde la cual observar al mundo, esto puede parecemos algo así como un rompecabezas. ¿Por qué no se producen estas rebeliones? En un organismo multicelular, cada célula tiene su propia tira de ADN, un completo conjunto de genes para construir un organismo en su totalidad, y si los genes son egoístas, ¿por qué los genes en las células del pulgar o en las células del ala del águila cooperan tan dócilmente con el resto de los genes? ¿No cuentan como genes las copias de ADN en los pulgares y en las alas? (¿Se les ha denegado el voto? ¿Porqué lo han tolerado?). Como han sugerido el biólogo David Sloan Wilson y el filósofo de la biología Elliot Sober (Wilson y Sober 1994), se puede aprender mucho sobre nuestros problemas sociales de deserción (por ejemplo, prometer y renegar de esta promesa) y de la tragedia de Hardin de las propiedades comunales (véase el capítulo 9) si se considera cómo nuestros ancestros, retrocediendo hasta los primeros eucariotas, trataban de conseguir «armonía y coordinación de sus partes». Sin embargo, las lecciones que han de aprenderse son engañosas, debido a que las células que nos constituyen pertenecen a dos categorías diferentes.

Un hombre medio es normalmente el anfitrión de miles de millones de organismos simbióticos pertenecientes quizás a miles de diferentes especies… Su fenotipo no está determinado solamente por sus genes humanos, sino también por los genes de todos los simbiontes que lo infectan. Las especies simbióticas que transporta un individuo tienen generalmente una muy variada procedencia, con solo unas cuantas procedentes de sus padres (Delius 1991:85).

¿Soy yo un organismo, o una comunidad, o ambas cosas? Soy ambas cosas —y más— pero hay una tremenda diferencia entre las células que son oficialmente parte de mi cuerpo y las células, muchas de ellas precisamente tan importantes para mi supervivencia, que no lo son. Las células que me componen como un yo multicelular, comparten todas un ancestro; son un linaje singular, las «células hijas» y las «células nietas» del óvulo y del

esperma que unidos formaron mi cigoto. Ellas son las células *huéspedes*; las otras células son *visitantes*, algunas bienvenidas y otras no. Las visitantes son forasteras, ya que descienden de linajes diferentes. ¿En qué consiste esta diferencia?

Es extremadamente fácil perder la objetividad, especialmente en contextos en los cuales tratamos todas estas «partes» como sistemas intencionales, como se debe, aunque con extrema precaución. A menos que seamos cuidadosos, podemos pasar por alto el hecho de que hay momentos cruciales en las «carreras» de estos varios agentes y semiagentes y hemi-semi-agentes cuando surgen oportunidades para «decidir» y entonces se dejan pasar. Las células que componen mi masa tienen un destino compartido, aunque algunas en un sentido más fuerte que otras. El ADN de las células de mi dedo y de las células de la sangre, se encuentran en un callejón sin salida genético; de acuerdo con la terminología de Weissmann (véase el cap. II) estas células son parte de la línea *somática* (el cuerpo), no de la línea *germinal* (las células sexuales). Exceptuando las revoluciones en las técnicas de clonación (e ignorando las perspectivas estrictamente limitadas y efímeras que tienen que dar origen a células para reemplazamiento que estas técnicas ayudan a crear) mis células de la línea somática están predestinadas a morir «sin hijos» y dado que esto fue determinado hace tiempo, no hay en modo alguno ni presión, ni oportunidad normal, ni «puntos de elección», como circunstancias en las cuales sus trayectorias intencionales —o las trayectorias de su limitada progenie— puedan ser corregidas. Se puede decir que son sistemas intencionales *balísticos*, cuyos más elevados objetivos y propósitos han sido fijados de una vez por todas, sin ninguna oportunidad de reconsideración o de guía. Son totalmente esclavos comprometidos con el *summum bonum* (‘el sumo bien’) del cuerpo del cual forman parte. Pueden ser explotadas o engañadas por los visitantes, pero bajo condiciones normales no pueden rebelarse contra su propio cuerpo. Como las mujeres de la película *Las esposas de Stepford*, tienen un único *summum bonum* diseñado entre ellas, y éste no es precisamente «ser el número uno». Son, por el contrario y por su propia naturaleza, jugadoras en equipo.

Como aquellas esposas, este *summum bonum* está diseñado también en ellas, y, desde este punto de vista, difieren fundamentalmente de las otras células que se encuentran «en el mismo barco», y que son mis simbióticos visitantes. Los benignos mutualistas, los comensales neutrales, y los deletéreos parásitos que comparten el vehículo que entre todos componen —es decir, yo—, tienen todos ellos su propio *summum bonum* diseñado en ellos

mismos, y posteriormente estará en sus respectivos linajes. Afortunadamente, existen condiciones bajo las cuales puede mantenerse una *entente cordiale*, ya que, al fin y al cabo, todas se encuentran en el mismo barco, y son limitadas las condiciones bajo las cuales obtendrían un mejor resultado sin cooperar. Pero son ellas las que han de hacer la elección. En cierto sentido, este es un problema de las células visitantes y no de las células huéspedes.

¿Por qué? ¿Qué hace posible —o qué empuja— a que las células huéspedes estén tan comprometidas, mientras que da rienda suelta a las células visitantes para rebelarse cuando surge la oportunidad? Ningún tipo de célula es, naturalmente, un agente racional, pensante y perceptor. Y ningún tipo de célula es significativamente más cognitivo que otro. No es aquí donde se localiza el punto de apoyo de la teoría del juego evolucionista. Los árboles secoyas no son señaladamente más listos, pero se encuentran en unas condiciones de competición que los fuerzan a abandonar, creando lo que es, desde *sus* puntos de vista (¡) una tragedia baldía. El mutuo acuerdo cooperativo por medio del cual todos renunciarían a crecer con troncos muy altos, y el abandono de sus vanos intentos para ganar más luz solar que la correspondiente a una razonable cuota, no se puede hacer cumplir desde el punto de vista de la evolución.

La condición que crea una elección es el insensato «voto» de la reproducción *diferencial*. Es la oportunidad de la reproducción diferencial la que permite a los linajes de nuestras visitantes «cambiar de ideas» o «reconsiderar» las elecciones que habían hecho, «explorando» políticas alternativas. Sin embargo, mis células huéspedes han sido diseñadas de una vez por todas, por un voto singular en el momento en que mi cigoto se formó. Si, gracias a una mutación, *ellas* adoptan estrategias dominantes o egoístas, entonces no prosperarán (en relación con sus contemporáneas) dado que hay una reducida oportunidad para la reproducción diferencial. (El cáncer puede ser considerado como una rebelión egoísta —y destructora del vehículo— hecha posible por una revisión que permite una reproducción diferencial).

El filósofo y lógico Brian Skyrms (1993, 1994a, 1994b) ha señalado recientemente que la precondición para la cooperación normal en el destino fuertemente compartido de las líneas somáticas de las células, es análoga a la cooperación que Rawls trata de construir tras el denominado velo de la ignorancia. Skyrms lo denomina, apropiadamente, el velo darwiniano de la ignorancia. Nuestras células sexuales (óvulo o esperma) están formadas mediante un proceso distinto de la división celular normal o *mitosis*. Nuestras células sexuales se forman mediante un proceso diferente, denominado

meiosis, el cual construye aleatoriamente la *mitad* del genoma candidato (para unir fuerzas con una mitad de la pareja) escogiendo, en primer lugar, un *bit* de la «columna A» (los genes procedentes de nuestra madre) y después un *bit* de la «columna B» (los genes procedentes de nuestro padre) hasta que una completa dotación de genes —aunque solo con una copia de cada uno— es construida e instalada en una célula sexual, lista para probar su suerte en la gran lotería del acoplamiento. Pero ¿qué «hijas» de nuestro original cigoto están destinadas para meiosis y cuáles para mitosis? Esto, también, es una lotería. Gracias a este mecanismo insensato, los genes paternos y maternos (en nosotros) no pueden «conocer su destino» por adelantado. La cuestión de si van a tener progenie de la línea germinal que pueda producir un flujo de descendientes que fluye hacia el futuro, o van a ser relegadas a las estériles aguas estancadas de la esclavitud de la línea somática, por el bien del cuerpo político o corporación (pienso en la etimología), es desconocida e irreconocible, así que no se gana nada con la competición egoísta entre los genes «compañeros».

Esta es, en cualquier caso, la disposición habitual. Sin embargo, hay ocasiones especiales, en las que el velo darwiniano de la ignorancia es alzado brevemente. Ya lo hemos comentado; están los casos del *impulso meiótico* o de la *impresión genómica* que consideramos en el capítulo 9 (Haig y Grafen 1991, Haig 1992), en cuyas circunstancias se *permite* una competición egoísta entre genes con el fin de elevarse; y elevarse quiere decir escalar la puesta en marcha de una carrera de armamento. Pero, en la mayoría de las circunstancias, para los genes «el tiempo en el que se puede ser egoísta» está estrictamente limitado y, una vez la suerte está echada, aquellos genes están de nuevo en marcha, hasta la próxima elección**[160]**.

Skyrms muestra que cuando los elementos individuales de un grupo —sea del organismo total o de sus partes— están estrechamente relacionados (clonas o casi clonas) o son, de otro modo, capaces de implicarse en el reconocimiento mutuo y en el «acoplamiento» ordenado, el simple modelo de la teoría del juego, conocido como el dilema del prisionero, en el cual siempre domina la estrategia de la traición, no puede utilizarse como modelo en aquellas circunstancias. Esto es así porque nuestras células somáticas no desertan; ellas son clonas. Esta es *una* de las condiciones bajo las cuales los grupos —tales como el grupo de mis células «huéspedes»— pueden tener la «armonía y la coordinación» requeridas para comportarse de una manera completamente estable, como un «organismo» o un «individuo». Pero antes de que gritemos tres hurras y aceptemos a éste como nuestro modelo para

hacer una sociedad justa, debemos hacer una pausa para señalar que hay otra forma de contemplar estos modelos de ciudadanos, las células y órganos de la línea somática; su particular tipo de abnegación es como la incuestionable obediencia de los fanáticos o de los zombies, mostrando la lealtad que corresponde a un grupo fieramente xenófobo, y es muy difícilmente un ideal para la emulación humana.

Nosotros, a diferencia de las células que nos componen, no estamos situados en trayectorias balísticas; somos misiles *dirigidos*, capaces de alterar el curso en cualquier punto de nuestra trayectoria, abandonando objetivos, cambiando lealtades, formando facciones y, después, traicionándolas, y así sucesivamente. Para nosotros, siempre es tiempo de decisión, y debido a que vivimos en un mundo de memes, ninguna consideración nos es ajena, o a una conclusión sacada de antemano. Por esta razón, estamos constantemente enfrentados con oportunidades sociales y dilemas del tipo de los que la teoría de los juegos aportan al campo de juego y las reglas de participación, pero no la solución. Cualquier teoría acerca del nacimiento de la ética tendrá que integrar cultura con biología. Como hemos dicho antes, la vida, para la gente que vive en sociedad, es más complicada.

2. *Friedrich Nietzsche y sus historias*

> El primer impulso para publicar algo de mi hipótesis sobre el origen de la moral me lo proporcionó un claro, pulcro y perspicaz —también precoz— librito en el cual encontré claramente, por primera vez, una especie de hipótesis genealógica, perversa y puesta patas arriba, el genuino tipo inglés, que me atrajo con ese poder de atracción que posee todo lo que nos es contrario, todo lo que está en nuestros antípodas.
>
> Friedrich Nietzsche,
> *La genealogía de la moral*

> Es en perfecto acuerdo con el esquema de la naturaleza, como a través de la selección natural, que la materia excretada para liberar el sistema de sustancias superfinas o nocivas deba ser utilizada para otros muy útiles propósitos.
>
> Charles Darwin,
> *On the Various Contrivances by Which Orchids Are Fertilised by Insects*

Friedrich Nietzsche publicó su *Genealogía de la moral* en 1887. Era el segundo gran sociobiólogo y, al contrario que Hobbes, fue inspirado (o provocado) por el darwinismo. Como he señalado en el capítulo 7, Nietzsche probablemente nunca leyó a Darwin. Su desdén para con el «tipo inglés» de

genealogía iba dirigido contra los darwinistas sociales: en particular, contra Herbert Spencer y los partidarios de Darwin en el continente europeo. Un partidario de Darwin era el amigo de Nietzsche, Paul Rée, cuyo «ordenado» libro titulado *Der Ursprung der moralischen Empfindungen* (1877) provocó la desordenada obra maestra de Nietzsche[161]. Los darwinistas sociales eran sociobiólogos, aunque ciertamente no de gran categoría. De hecho, sus esfuerzos se limitaron a los memes de su héroe, popularizando (per)versiones de Darwin de segunda categoría.

La «supervivencia del mejor adaptado», proclamaba Spencer, no es precisamente el camino de la madre naturaleza, sino que *debe ser nuestro* camino. Según los darwinistas sociales, es «natural» que el fuerte venza al débil, y que el rico explote al pobre. Esto es simplemente una reflexión deficiente, y Hobbes ya nos había mostrado el porqué. Es igualmente «natural» morir joven y analfabeto, sin beneficiarse de gafas para la miopía o medicina para la enfermedad —porque es así como era en el estado natural— pero es seguro que esto no sirve para nada cuando preguntamos: ¿debe ser éste ahora nuestro camino? De modo alternativo, como para nosotros era (en un sentido amplio) enteramente natural —y no sobrenatural— salir del estado de naturaleza y adoptar un sinnúmero de prácticas sociales en nuestro mutuo beneficio, podemos simplemente negar que haya algo universalmente natural en que el fuerte domine al débil, y para el resto del darwinismo social es un desatino. Es divertido señalar que el fundamental (mal) argumento del darwinismo social es idéntico al (mal) argumento utilizado por muchos fundamentalistas religiosos. Mientras que los fundamentalistas comienzan, a veces, su argumentación diciendo «Si Dios ha creado al hombre para… (volar, usar vestidos, beber alcohol…)», los darwinistas sociales comienzan las suyas diciendo, en efecto, «Si la madre naturaleza ha querido hacer al hombre para…», e incluso aunque la madre naturaleza (la selección natural) puede ser considerada como poseedora de intenciones, en el sentido limitado de haber apoyado retrospectivamente hechos, por una razón u otra, estos precoces apoyos no cuentan ahora para nada, ya que las circunstancias han cambiado.

Entre las ideas de los darwinistas sociales había una agenda política: los esfuerzos de «pretendidos benefactores» para proporcionar alimentación y educación a los miembros menos afortunados de la sociedad son contraproducentes; estos esfuerzos permiten reproducirse a aquellos a quienes la naturaleza sabiamente desecha. Estas son ideas abominables, pero no eran el primer objetivo de las críticas de Nietzsche. Su primer objetivo era la

histórica ingenuidad de los darwinistas sociales (Hoy 1986), su optimismo panglossiano acerca de la fácil adaptabilidad de la razón humana (o prudencia de Hobbes) a la moral. Nietzsche interpretó esta complacencia de los darwinistas sociales como parte de su herencia como «psicologistas ingleses», descendientes intelectuales de Hume. Nietzsche hace notar su deseo de evitar los ganchos celestes:

> ¿Qué desean realmente los psicologistas ingleses? Uno siempre los descubre… buscando el agente verdaderamente efectivo y directo, que ha sido decisivo en su evolución, en ese lugar donde el orgullo intelectual del hombre menos *desea* encontrarlo (en la *vis inertiae* del hábito, por ejemplo, o en su descuido, o en un conglomerado de ideas ciegas y mecanicistas y ligadas al azar, o en algo puramente pasivo, automático, reflexivo, molecular y completamente estúpido), ¿qué es realmente lo que siempre lleva a estos psicologistas precisamente por *esta* dirección? ¿Es un secreto malicioso, vulgar, quizás un instinto autoengañoso para hombres menospreciables? (*La genealogía de la moral*, I, §1).

El antídoto de Nietzsche contra las banalidades de los «psicologistas ingleses» era un romanticismo muy «continental». Pensaban que el paso del estado de naturaleza a la moral era fácil, o al menos completamente presentable, pero era porque ellos habían inventado sus historias y no se preocuparon de examinar las claves de la historia, cuyo relato es más misterioso.

Nietzsche comenzó, como hizo Hobbes, imaginando un mundo premoral de la vida humana, aunque dividió la historia de la transición en dos fases (y relató sus historias en orden inverso, comenzado por el medio, algo que confunde a muchos lectores). Hobbes (*Leviatán*, I, 14) había observado que la propia existencia de cualquier práctica de establecer contratos o convenios depende de la capacidad de los seres humanos para hacer promesas acerca del futuro, y lo que llamaba la atención de Nietzsche era que esta capacidad no llegara gratuitamente. Este era el tema del segundo ensayo de los tres que componen *La genealogía de la moral*: «Criar a un animal *con el derecho a hacer promesas*, ¿no es la paradójica tarea que la naturaleza se ha impuesto en el caso del hombre?, ¿no es el problema real en lo que se refiere al hombre?» (II, § 1). Esta «larga historia acerca de cómo se ha originado la *responsabilidad*» es la historia de cómo los primeros seres humanos aprendieron a torturarse entre ellos —literalmente— mediante el desarrollo de un tipo especial de memoria, la necesaria para guardar las cuentas del debe y del haber. «Comprar y vender, junto con sus adimentos psicológicos, son aún más viejos que los inicios de cualquier tipo de formas sociales de organización y de alianzas» (*ibidem*, II, §8). La capacidad de detectar trampas, de recordar las promesas rotas y de castigar al tramposo hubo de ser inculcada en los cerebros de nuestros ancestros; Nietzsche resume: «Sus

comienzos estuvieron, como los comienzos de todo lo magnífico sobre la tierra, completamente empapados en sangre, y durante largo tiempo» (*ibidem*, II, §6). ¿Cuál es la evidencia de Nietzsche acerca de todo esto? Una lectura imaginativa —por no decir desmedida— de lo que podían ser denominados los archivos de fósiles de la cultura humana, bajo la forma de antiguos mitos, prácticas religiosas que sobreviven, claves arqueológicas, etc. Dejando a un lado los detalles cruentos, aunque sean fascinantes, la sugerencia de Nietzsche es que eventualmente —quizás a través de un ejemplo del llamado efecto Baldwin— nuestros ancestros «criaron» un animal con una innata capacidad para mantener una promesa y un concomitante talento para detectar y castigar a los que no cumplieran lo prometido.

Esto permitió, según Nietzsche, la formación de las primeras sociedades, pero aún no existía moralidad, al menos en el sentido en que hoy la reconocemos y honramos. La segunda transición, afirma Nietzsche, ocurrió en tiempos históricos y puede ser seguida a través de la reconstrucción etimológica y una lectura apropiada de los textos de los dos últimos milenios, mediante una adaptación realizada por Nietzsche de los métodos filosóficos en cuyo uso había sido entrenado. Para leer estas claves de un modo nuevo se necesita una teoría, naturalmente, y Nietzsche la tiene, desarrollada en oposición a la tácita teoría que según él usaban los darwinistas sociales. Los protociudadanos de la segunda «historia al fin y al cabo», al estilo de Kipling, relatada por Nietzsche (*ibidem*, I) viven en sociedades de esta suerte, no en el estado natural de Hobbes, aunque la vida que se describe en ellas es, del mismo modo, detestable y brutal. El poder se convierte en lo correcto o, más bien, el poder hace las reglas. La gente posee los conceptos de lo bueno y de lo malo, pero no del bien y del *mal* lo correcto y lo erróneo. Al igual que Hobbes, Nietzsche trata de contar la historia de cómo surgieron estos últimos memes. Una de las más osadas (y, en ultimo término, menos persuasiva) de sus especulaciones es que los memes para el bien (moral) y el mal no fueron permutaciones menores de sus predecesores amorales; los memes *intercambiaron sus posiciones*. Lo que había sido *bueno* (al viejo estilo) llegó a ser el *mal* (al nuevo estilo) y lo que había sido *malo* (al viejo estilo) llegó a ser (moralmente) *bueno* (al nuevo estilo). Este «cambio de valores» era para Nietzsche el acontecimiento clave en el nacimiento de la ética y él se opone explícitamente a la blanda suposición de Herbert Spencer de que

el concepto «bueno» es esencialmente idéntico al concepto «útil» y «práctico», así que en los juicios «bueno» y «malo» la humanidad ha resumido y sancionado precisamente sus *inolvidadas e inolvidables* experiencias respecto a lo que es útil-práctico y lo que es dañino-impráctico. De acuerdo con esta teoría, lo que siempre se ha comprobado útil en sí mismo es bueno; por consiguiente se podría afirmar que es

«valioso en el grado más elevado», ya que es «valioso en sí mismo». Este camino hacia una explicación es, como ya vimos, también erróneo, aunque, al menos, la explicación es, en sí misma, razonable y psicológicamente sostenible (*ibidem*, I, §3).

El maravilloso e ingenioso relato de Nietzsche nos cuenta cómo el intercambio de valores que tuvo lugar hace difícil un resumen razonable y, a menudo, es presentado de manera terriblemente incorrecta. No tengo que hacerle justicia en este momento, pero desearía llamar la atención hacia su tema central (sin juzgar si es verdad): los «aristócratas» que establecieron las reglas gracias a su poder sobre los débiles fueron astutamente engañados (por los «sacerdotes») al adoptar la inversión de los valores, y esta «revuelta de los esclavos en la moral» dirigió la crueldad del fuerte contra él mismo, de modo que el fuerte fue manipulado para que se avasallaran y civilizaran entre ellos mismos

Pues con los sacerdotes *todo* llegó a ser más peligroso, no sólo curas y remedios, sino también arrogancia, venganza, crítica, libertinaje, amor, pasión por el poder, virtud, enfernedad —aunque es honesto añadir que fue en la tierra de cultivo de esta forma de la existencia humana *esencialmente peligrosa* que es la forma sacerdotal, donde el hombre llegó a ser, por vez primera, un *animal interesante*, y que sólo aquí el alma humana, en su sentido más elevado, adquirió *hondura* y llegó el *mal* — y ¡estos son los dos aspectos básicos aspectos en los cuales el hombre ha sido hasta ahora superior a las otras bestias!

Las historias de Nietzsche, como las «historias al fin y al cabo» de Kipling, son fantásticas (al viejo y al nuevo estilo). Son una mezcla de una historia llena de talento y de locura, sublime e innoble, intensamente devastadora, con una fantasía sin límites. Si la imaginación de Darwin estaba hasta cierto punto menoscabada por su herencia mercantil inglesa, la de Nietzsche lo estaban aún más por su herencia intelectual germana, pero tales datos biográficos (cualesquiera que sean) no tienen relación con el valor actual de los memes, a cuyo nacimiento cada uno asistió tan brillantemente. Ambos llegaron con ideas peligrosas —si no me equivoco, esto no es coincidencia—, pero mientras que Darwin era extremadamente cauto en sus expresiones, Nietzsche se dejaba llevar de una prosa tan inflamada que sin duda sirvió para que su legión de devotos incluyera una nada reputable manada de incalificables e incomprensibles nazis y otros partidarios de esta ralea cuyas perversiones de sus memes hacen parecer la perversión de Darwin a manos de Spencer casi inocente. En ambos casos, debemos trabajar en la reparación del daño que tales descendientes han producido en nuestros propios filtros de memes, los cuales tienden a descartar memes sobre la base de la culpabilidad por

asociación. Ni Darwin ni Nietzsche eran políticamente correctos, afortunadamente para nosotros.

(Lo políticamente correcto, en las versiones extremas merecedoras de este nombre, es antitético con casi todos los sorprendentes avances del pensamiento. Podemos llamarlo *euménica*, dado que es, como la eugenesia extrema de los darwinistas sociales, un intento de imponer miópicamente estándares derivados de la seguridad y de la bondad sobre la liberalidad de la naturaleza. Pocos hoy —aunque hay unos cuantos— calificarían *todas* las consultas genéticas, todas las políticas genéticas con el título condenatorio de eugenesia. Debemos reservar este término crítico para las políticas insaciables y perentorias, las políticas extremistas. En el capítulo 18, consideraremos de qué manera podemos recorrer más inteligentemente la esfera de los memes y lo que podemos hacer para protegernos a nosotros mismos de las ideas verdaderamente peligrosas, aunque cuando lo hagamos no debemos olvidar el mal ejemplo de la eugenesia).

La contribución más importante de Nietzsche a la sociobiología ha sido, en mi opinión, la firme aplicación de una de las fundamentales intuiciones del propio Darwin en el campo de la evolución cultural. Esta es la intuición darwiniana más notoriamente pasada por alto por los darwinistas sociales y por algunos sociobiólogos contemporáneos. Este error ha sido denominado a veces como la «falacia genética» (por ejemplo. Hoy 1986): el error de inferir una función o significación actual de una ancestral función o significación. Como Darwin puntualizó: «Así, a lo largo de la naturaleza casi cada parte de cada cosa viva ha servido probablemente, en una condición ligeramente modificada, para diversos propósitos, y ha actuado en la maquinaria viviente de muchas antiguas y distintas formas específicas» (1862:284). Y Nietzsche escribió:

> La causa del origen de una cosa y su eventual utilidad, su lugar y su empleo real en un sistema de propósitos, se encuentra en un mundo aparte; donde quiera que exista, habiendo llegado a ser de alguna manera, es reinterpretada para nuevos fines una y otra vez, controlado, transformado y redirigido por algún poder superior; todos los eventos en el mundo orgánico son un sometimiento, un *llegar a ser el que manda*, y todo el proceso de someter y llegar a ser el que manda implica una nueva interpretación, una adaptación a través de la cual cualquier «significado» y «propósito» previos son necesariamente oscurecidos o incluso obliterados (*La genealogía de la moral*, II, §12).

Aparte de la indignación característica de Nietzsche sobre el poder que somete y llega a ser el amo, esto es puro Darwin. O, como Gould puede decir, todas las adaptaciones son exadaptaciones tanto en la evolución cultural como en la evolución biológica. Nietzsche también hizo hincapié en otro clásico tema darwiniano:

La «evolución» de una cosa, una costumbre, un órgano, no ha de entenderse, bajo ningún concepto, como su *progreso* hacia un objetivo, incluso menos aún como un *progreso* lógico por el camino más corto y con el menor gasto de fuerza, sino como una sucesión de procesos de sometimiento más o menos profundos, más o menos mutuamente independientes, más la resistencia que estos procesos puedan encontrar, los intentos de transformación con el propósito de defensa y reacción, y los resultados de contraataques exitosos Nietzsche (*ibidem*, II, §12)**[162]**.

Teniendo en cuenta que Nietzsche nunca había leído la obra de Darwin, su apreciación de sus ideas fundamentales es notable, pero, al poco, estropea su *curriculum* como darwiniano con fundamento, al caer, en la misma página, en la pasión por los ganchos celestes, anunciando su «fundamental oposición al nuevo prevalente instinto y gusto que sería mejor reconciliar con lo absolutamente fortuito, lo mecánico y sin sentido de todos los acontecimientos que con la teoría que postula que en todos los acontecimientos opera una *voluntad de poder*». La idea de Nietzsche de la voluntad de poder es una de las más extrañas encarnaciones de la pasión por los ganchos celestes y, afortunadamente, pocos la encuentran hoy atractiva. Pero, si la dejamos a un lado, el resultado de la genealogía de la moral de Nietzsche es que debemos ser extremadamente cuidadosos para no leer en la historia que extrapolamos de la naturaleza ninguna conclusión simplista acerca del valor:

La cuestión: ¿cuál es el *valor* de ésta o de esa tabla de valores morales?, debe ser contemplada desde las más diversas perspectivas: pues el problema «¿valor para *qué*?» no puede ser examinado con demasiada sutileza. Algo, por ejemplo, que posee un valor obvio en relación con la más larga supervivencia posible de una raza (o para el reforzamiento de su poder de adaptación a un clima concreto o para la preservación del mayor número) bajo ningún concepto posee el mismo valor si fuera la cuestión, por ejemplo, de producir un tipo más fuerte. El bienestar de la mayoría y el bienestar de unos pocos son puntos de vista opuestos sobre el valor: considerar el primero, *a priori*, de valor más alto podemos dejarlo a la ingenuidad de los biólogos ingleses (*La genealogía de la moral*, I, §17).

Es a Spencer, sin duda, y no a Darwin, a quien Nietzsche acusa de ingenuidad acerca del valor moral. Tanto Spencer como Rée pensaron que podían observar una vía directa y simple hacia el altruismo (Hoy 1986:29). Cabe considerar las críticas de Nietzsche a este panglossianismo como un claro adelantado de la críticas de George Williams al panglossianismo del ingenuo grupo seleccionista (véase el capítulo 11). Spencer, según nuestra terminología, fue un egregio reduccionista insaciable, que trataba de derivar el «debe» del «es» en un solo paso. Pero, ¿no revela esto el más profundo problema de toda la sociobiología? ¿No han mostrado los filósofos que nunca podemos derivar «debe» de «es», no importa cuántos pasos se den? ¡Algunos han argumentado que la sociobiología, no importa lo avanzada que llegue a

ser, no importa cuántas grúas utilice, nunca podrá tender un puente en el hiato que existe entre el «es», el hecho empírico científico, y el «debe» de la ética! (Dicen esto con una pasión impresionante). Es esta convicción apasionada la que vamos a examinar a continuación.

3. *Algunas variedades del reduccionismo ético insaciable*

Uno de los lemas característicos de la filosofía contemporánea es que no podemos derivar «debe» de «es». El intento de hacer esta derivación se denomina a menudo la *falacia naturalista*, término procedente de la obra clásica de G. E. Moore *Principia Ethica* (1903). Como señala el filósofo Bernard Williams (1983:5 56), hay realmente varias cuestiones en este punto. El naturalismo «consiste en el intento de establecer ciertos aspectos fundamentales de la vida buena para el hombre sobre la base de consideraciones de la naturaleza humana». El naturalismo no podría ser refutado por el hecho obvio de que no podemos derivar ninguna simple declaración de «debe» de una simple declaración de «es». Veamos: ¿se sigue lógicamente que yo *debo* darle cinco dólares por el hecho (y supongamos que es un hecho) de que yo dijera que le daría cinco dólares? Obviamente no; se pueden alegar muchas condiciones justificativas que intervienen para bloquear esta inferencia. Incluso si llegáramos a caracterizar lo que mencioné como una *promesa* —una descripción cargada éticamente—, de ello no se infiere directamente una *simple* declaración de «debe».

Pero reflexiones como ésta hacen escasa mella en el naturalismo como objetivo teórico. Los filósofos distinguen entre lo que es encontrar las condiciones *necesarias* y las condiciones *suficientes* para varias cosas, y la aplicación de esta distinción en este caso ayuda realmente a clarificar la situación. Una cosa es negar que colecciones de hechos acerca del mundo natural sean *necesarias* para fundamentar una conclusión ética, y otra cosa completamente distinta es negar que cualquier colección de tales hechos sea *suficiente*. Según la doctrina estándar, si permanecemos firmemente asentados en el reino de los hechos acerca de cómo es el mundo, nunca encontraremos colección alguna de estos hechos, tomados como axiomas, a partir de la cual *pueda ser probada de forma concluyente* ninguna conclusión ética específica. No podemos pasar de la una a la otra como tampoco podemos deducir todas las proposiciones verdaderas de la aritmética a partir de un conjunto coherente de axiomas de aritmética.

Bien, ¿y qué? Podemos sacar a relucir la fuerza de esta cuestión retórica con otra, algo más directa: si «debe» no puede derivarse de «es», ¿de dónde *puede* derivarse «debe»? ¿Es la ética un campo de investigación *enteramente* autónomo? ¿Flota a su aire, desligada de los hechos de cualquier otra disciplina o tradición? ¿Surgen nuestras intuiciones morales de algún inexplicable módulo ético implantado en nuestros cerebros (o en nuestros «corazones», para hablar de acuerdo con la tradición)? Este sería un dudoso gancho celeste en el cual se colgarían nuestras más profundas convicciones acerca de lo que es correcto y lo que es erróneo. Colin McGinn escribe:

> Según Chomsky, es plausible considerar a nuestra facultad ética como análoga a nuestra facultad del lenguaje; adquirimos conocimiento ético con una instrucción muy poco explícita, sin gran trabajo intelectual y al final el resultado es notablemente uniforme dada la variedad del *input* ético recibido. El medio ambiente sirve simplemente para poner en marcha y especializar un esquematismo innato… En el modelo de Chomsky, tanto la ciencia como la ética son productos naturales de la contingente psicología humana, limitada por sus principios específicos constitutivos; sin embargo, la ética parece tener una base más segura en nuestra arquitectura cognitiva. Hay un elemento de azar en nuestra posesión del conocimiento científico que no existe en el caso de nuestro conocimiento ético (Me Ginn 1993:30).

Ante el contraste de nuestro presunto sentido innato del conocimiento ético con nuestra simplemente «azarosa» capacidad para ocuparnos de la ciencia, McGinn y Chomsky sugieren que hay *razones* para nuestra posesión de la ética que han de ser descubiertas. Si hubiera un módulo moral, ciertamente desearíamos conocer en qué consiste, cómo ha evolucionado, y lo más importante de todo, por qué existe. Pero, de nuevo, si intentamos penetrar en el problema. Me Ginn trata de cerrar la puerta atrapándonos los dedos, al desacreditar con el calificativo de «cientifismo» el intento de dar respuestas a nuestros interrogantes científicos respecto a la fuente de esta maravillosa perspectiva que poseemos nosotros, y no las otras criaturas.

¿De dónde puede derivarse este «debe»? La respuesta más atrayente es ésta: la ética debe basarse, de *alguna manera*, en nuestra apreciación de la naturaleza humana, en el sentido de lo que un ser humano es o puede ser, y en lo que un ser humano puede desear tener o desea ser. Si *esto es naturalismo*, entonces el naturalismo no es una falacia. Nadie puede negar seriamente que la ética es responsable de tales hechos en relación con la naturaleza humana. Podemos estar en desacuerdo sobre dónde hemos de buscar los hechos más significativos de la naturaleza humana: en novelas, en textos religiosos, en experimentos psicológicos, en investigaciones biológicas y antropológicas. La falacia no es naturalismo, sino más bien cualquier intento mentalmente simplista de pasar rápidamente de los hechos a los valores. En otras palabras,

la falacia es un reduccionismo insaciable que pasa de los valores a hechos, más que un reduccionismo considerado de una manera más cautelosa, como el intento de unificar nuestra visión del mundo, de modo que nuestros principios éticos no choquen irracionalmente con el modo según el cual el mundo *es*.

La mayoría de los debates acerca de la falacia naturalista es mejor interpretarlos como desacuerdos análogos a los debates entre ganchos celestes y grúas en la teoría la evolución. Por ejemplo, B. E Skinner, en mi opinión el paladín de todos los tiempos en el reduccionismo insaciable, escribió un tratado de ética titulado *Más allá de la libertad y la dignidad* (1971). En este tratado, Skinner «se compromete con la falacia naturalista» a todos los ni veles, desde lo diminuto a lo megalomaníaco. «Hacer un juicio de valor llamando a algo bueno o malo es clasificarlo en términos de sus efectos reforzantes» (Skinner 1971:105). Veamos: ¿significaría esto que la heroína es buena, aparentemente, y que cuidar de los padres ancianos es malo? ¿Es esta objeción ser quisquilloso con una definición nada cuidadosa? El efecto reforzador de la heroína, nos asegura Skinner cuando cae en la cuenta del problema (p. 110), es «anómalo». Esta respuesta difícilmente puede ser una defensa convincente contra la acusación de reduccionismo insaciable. Skinner insiste en su libro sobre el carácter científico de «su diseño para una cultura», y cómo es muy apropiado para… ¿para qué? ¿Cuál es su caracterización del *summum bonum*?

Nuestra cultura ha producido la ciencia y la tecnología que necesita para salvarse a sí misma. Posee la riqueza necesaria para una acción efectiva. Demuestra, hasta un punto considerable, una preocupación por su propio futuro. Pero si continúa apostando por la libertad y la dignidad, más que por su propia supervivencia, como su principal valor, entonces es posible que alguna otra cultura pueda hacer una mayor contribución en el futuro (Skinner 1971:181).

Espero que el lector comparta conmigo el rechazo a esta reflexión de Skinner: ¿Qué es esto? Incluso si Skinner estuviera en lo cierto al pensar (y seguramente no lo está) que un régimen conductista es la mejor opción para preservar nuestra cultura para el futuro, espero que quede claro para todos mis lectores que Skinner pudo haberse equivocado cuando creyó que la «supervivencia de la cultura» era el objetivo más elevado que cualquiera de nosotros pudiera incluso imaginar como deseable. En el capítulo 11 hemos considerado brevemente lo malo que sería anteponer la supervivencia de nuestros propios genes a todas las cosas. ¿Es la supervivencia de nuestra propia cultura un objeto claramente más razonable, para colocarlo en un pedestal por encima de todas las cosas? ¿Justificaría esta postura los crímenes

en masa, por ejemplo, o traicionar a todos nuestros amigos? Nosotros, utilizadores de memes, podemos entrever otras posibilidades, más allá de nuestros genes y más allá, incluso, del bienestar de los grupos y culturas a los cuales actualmente pertenecemos. Contrariamente a lo que le sucede a nuestras células de la línea somática, somos capaces de concebir más complicadas *raisons d'être*.

¡El error de Skinner no es que intentara fundamentar la ética en hechos científicos acerca de la naturaleza humana, sino que sus intentos fueran demasiado simplistas! Yo supongo que las palomas podrían prosperar tan bien como pudieran desear en la utopía skinneriana, pero nosotros somos realmente bastante más complicados que los palomas. El mismo defecto se observa en la incursión en la ética de otro profesor de Harvard, E. O. Wilson [1975], uno de los más grandes entomólogos del mundo y el acuñador del término «sociobiología». En su tratado ético titulado *Sobre la naturaleza humana* (1978), Wilson encara el problema de identificar el *summum bonum* o «valor cardinal» y plantea la existencia de dos valores recíprocos: «Los cultivadores de la nueva ética desearían ponderar en los comienzos el valor cardinal de la supervivencia de los genes humanos, bajo la forma de un reservorio común a lo largo de las generaciones… Yo creo que una correcta aplicación de la teoría de la evolución favorece también la diversidad en el reservorio de genes como un valor cardinal» (1978:196-198). Wilson añade un tercer valor cardinal, los derechos humanos universales, aunque sugiere que debe ser desmitificado. Una «hormiga racional» encontraría que el ideal de los derechos humanos «no tiene fundamentos biológicos y que el mismo concepto de la libertad individual es intrínsecamente perverso» (*ibidem*: 199).

> Deseamos acceder a los derechos universales porque en las sociedades tecnológicas avanzadas el poder es demasiado fluido para evitar este imperativo de los mamíferos; las consecuencias a largo plazo de la desigualdad serán siempre visiblemente peligrosas para sus beneficiarios temporales. Sugiero que esta es la verdadera razón del movimiento de los derechos humanos y que una comprensión de esta cruda causa biológica será al final más atrayente que cualquier racionalización inventada por la cultura para reforzar a estos derechos y expresarlos con eufemismos (E. O. Wilson 1978:199).

Escribiendo en colaboración con el filósofo de la biología Michael Ruse, Wilson declara que la sociobiología nos ha demostrado que la «moral o, más exactamente, nuestra creencia en la moral, es simplemente una adaptación introducida para facilitar nuestros fines reproductivos» (Ruse y Wilson 1985). Una insensatez. Nuestros fines reproductivos pueden haber sido los fines que nos han mantenido funcionando hasta que hemos podido desarrollar la cultura y pueden aun desempeñar un papel poderoso —a veces excesivamente

poderoso— en nuestro pensamiento, pero esto no autoriza conclusión alguna acerca de nuestros valores actuales. Del hecho de que nuestros fines reproductivos fueran la última *fuente* histórica de nuestros valores presentes no se sigue que sean los últimos (y todavía principales) *beneficiarios* de nuestras acciones éticas. Si Ruse y Wilson piensan de otra manera, están cometiendo la falacia «genética» de la que Nietzsche (y Darwin) nos previnieron. Como dice Nietzsche, «la causa del origen de una cosa y su eventual utilidad, su empleo en la realidad y su lugar en un sistema de propósitos, es un mundo aparte». ¿Han cometido Ruse y Wilson esta falacia? Veamos lo que dicen sobre este tema:

Es un importante sentido, la ética como nosotros la comprendemos es una ilusión que nos ha sido inoculada por nuestros genes para inducirnos a cooperar… Además, el camino con el que nuestra biología refuerza sus fines es haciéndonos creer que hay un objetivo de un código más elevado, al cual todos estamos sujetos (Ruse y Wilson 1985:51).

Debe ser verdad que hay una explicación evolucionista sobre cómo nuestros memes y nuestros genes interactúan para crear las políticas de la cooperación humana que disfrutamos en la civilización —no hemos calculado todos los detalles todavía, pero debe ser verdad, a menos que haya ganchos celestes en la lejanía— pero esto no demostraría que el resultado fuera *para el beneficio de los genes* (como principales beneficiarios). Una vez que los memes aparecen en escena, ellos, y las *personas* que ayudan a crear, son también beneficiarios potenciales. En consecuencia, la verdad de una explicación evolucionista no demostraría que nuestra lealtad a los principios éticos o a un «código más elevado», fueran una «ilusión». En una famosa imagen, Wilson expresó así su visión del problema:

Los genes mantienen a la cultura atada con una correa y un collar. La correa es muy larga, aunque inevitablemente los valores estarán limitados de acuerdo con sus efectos sobre el reservorio de genes humanos (E. Wilson 1978:167).

Pero lo que todo esto significa (a menos que sea falso) es que, a largo plazo, si adoptamos prácticas culturales que tengan efectos desastrosos sobre el reservorio de genes humanos, entonces este reservorio sucumbirá. Sin embargo, no hay razón para pensar que la biología evolucionista nos demuestre que nuestros genes son lo bastante poderosos y perspicaces para impedirnos poner en marcha políticas completamente contrarias a sus intereses. Por el contrario, el pensamiento evolucionista nos muestra que nuestros genes pueden ser, a duras penas, más inteligentes que los ingenieros que diseñaron nuestras imaginarias máquinas de supervivencia (véase el

capítulo 14); ¡recordemos la impotencia de estos ingenieros ante la perspectiva de colaboraciones no previstas con otros robots! Hemos visto ejemplos de parásitos —como los virus— que manipulan la conducta de sus huéspedes para *sus* intereses, en lugar de los intereses de los huéspedes. Y hemos visto ejemplos de comensales y mutualistas que se unen en una causa común, para crear un beneficiario mayor surgido de las partes. Las personas, según el modelo del meme que hemos esquematizado, son precisamente entidades mucho más grandes y más elevadas, y las políticas que *ellas* llegan a adoptar como resultado de interacciones entre sus cerebros infectados de memes no están en absoluto obligadas a responder solamente a los intereses de sus genes o de sus memes. Esta es nuestra trascendencia, nuestra capacidad para «rebelarnos contra la tiranía de los replicadores egoístas», como dijo Dawkins, y no hay nada en esta afirmación que sea antidarwiniano o anticientífico.

La típica incapacidad de Wilson y otros sociobiólogos para considerar a sus críticos como algo más que fanáticos religiosos o esotéricos científicamente analfabetos, es, una vez más, una triste y excesiva oscilación del péndulo. Skinner calificaba a sus críticos como un hatajo de cartesianos dualistas y adoradores de los milagros, y en su perorata declara:

> Al hombre *qua* hombre, no dudamos en decirle ¡en buena hora me libré! Sólo desposeyéndolo a él podemos pasar a considerar las causas reales de la conducta humana. Sólo entonces podemos abandonar la inferencia en aras de la observación y pasar de lo milagroso a lo natural, de lo inaccesible a lo manipulable. (Skinner 1971:201).

Wilson y otros muchos sociobiólogos tienen la misma mala costumbre de considerar a cualquiera que no esté de acuerdo con ellos como un buscador de ganchos celestes, ignorante y temeroso de la ciencia. De hecho, ¡sólo la *mayoría* de la gente que no está de acuerdo con ellos se adapta a esta descripción! Hay una minoría que comprende a los críticos responsables de los excesos de lo insaciable, ante la cual los entusiastas exponentes de cualquier nueva escuela científica están en condiciones de sucumbir.

Otro eminente biólogo, Richard Alexander, cuyo tratamiento de la ética es bastante más cuidadoso, expresa el apropiado escepticismo sobre los valores cardinales propuestos por Wilson: «Se juzguen estos objetivos admirables o no para la humanidad, Wilson no relaciona la selección de estos valores con principios biológicos» (Alexander 1987:167). Pero Alexander también subestima el poder de la cultura —memes— para romper la correa de Wilson. Como Wilson, Alexander reconoce la enorme diferencia, en cuanto a velocidad, entre la evolución cultural y la genética, y alega con energía que la

versatilidad cultural deja maltrecho cualquier intento —como los de Chomsky y Fodor— de encontrar un límite, un «prohibido el paso» para el conocimiento humano. Piensa Alexander, sin embargo, que la biología evolucionista ha demostrado que «los propios intereses de los individuos no pueden realizarse únicamente a través de la reproducción, creando descendientes y asistiendo a otros parientes» y que una consecuencia de esto es que ninguno expresa beneficencia genuina o altruismo. Como él señala:

> La «más grande revolución intelectual del siglo» nos dice que, a pesar de nuestras intuiciones, no hay prueba alguna que apoye la tesis de la beneficencia, y una gran cantidad de teoría convincente sugiere que tal visión del mundo será, con el tiempo, juzgada como falsa (Alexander 1987:3).

Pero Alexander, como Wilson y los darwinistas sociales, se compromete con una sutil y atenuada versión de la falacia genética, y hace hincapié en el propio pasaje en el cual acepta este compromiso:

> Aunque la cultura cambiase masivamente y continuamente a lo largo de múltiples generaciones, aunque nuestros problemas y promesas surgiesen fuera del proceso cultural, aun cuando no hubiera variaciones genéticas entre los humanos que afectasen significativamente a su conducta, siempre será verdad que la historia acumulativa de la selección natural continúa influyendo en nuestras acciones por el conjunto de genes que ha aportado la humanidad (Alexander 1987:23).

Esto es naturalmente verdad, pero no confirma la idea que él mismo propone. Como Alexander insiste, sin importar cuán potentes sean las fuerzas culturales, éstas siempre han de actuar sobre los materiales que las fuerzas genéticas han conformado y seguirán conformando, aunque tales fuerzas culturales puedan igualmente *redirigir* o *explotar* o *subvertir* aquellos diseños genéticamente transferidos como *atenuarlos* o *combatirlos*. A los sociobiólogos, que reaccionan en exceso frente a los absolutistas culturales (aquellos locos de los ganchos celestes) de un modo semejante a como Darwin reaccionó frente a los catastrofistas, les gusta recalcar que la cultura debe haber *crecido a partir* de nuestra herencia biológica. Naturalmente esto es así, y es también verdad que nos hemos desarrollado a partir de los peces, pero nuestras razones no son las razones de los peces precisamente porque los peces sean nuestros ancestros.

Los sociobiólogos están en lo cierto cuando resaltan que nuestra única capacidad para adoptar y actuar de acuerdo con un grupo de razones diferente, no impide que seamos molestados o incluso torturados o traicionados por nuestras «urgencias animales». Mucho tiempo antes de que Salomé bailara la danza de los siete velos, era ya obvio para los miembros de nuestra especie que las innatas urgencias procreadoras pueden hacer su aparición en los

momentos más inoportunos, del mismo modo que estornudar o toser, amenazando seriamente el bienestar del cuerpo donde la urgencia aparece. Como sucede en otras especies, muchas son las mujeres que han perecido por salvar a su hijo, y muchos los hombres que han muerto prematuramente al proseguir vehementemente, por uno u otro peligroso rumbo, el impulso marcado por la débil esperanza de la procreación. Pero no debemos convertir este importante hecho acerca de nuestras limitaciones biológicas en la idea totalmente equivocada de que el imperativo de nuestros genes es el *summum bonum* que se halla el origen de toda secuencia de razonamientos prácticos. Un contraejemplo nos muestra por qué no lo es: Larry, enfermo de amor tras ser rechazado con desdén por Lola, el amor de su vida, se inscribe en el Ejercito de Salvación para tratar de olvidarla y poner fin a su tormento. Funciona. Años más tarde San Larry el Sublimado gana el Premio Nobel de la Paz por todas sus buenas obras y Richard Alexander, en la ceremonia de Oslo, echa un jarro de agua fría en el acto al recordarnos que todo esto tuvo su origen en el instinto reproductivo de Larry. Así fue. ¿Y qué? Cometeremos un gran error si creemos que para comprender la magnitud de la vida de Larry tenemos que interpretar todos y cada uno de sus movimientos como diseñados, por una vía indirecta u otra, para asegurar que tuviera nietos.

La posibilidad de que un meme o un complejo de mentes pueda cambiar la dirección de nuestras subyacentes proclividades genéticas es ilustrada llamativamente por un experimento sociobiológico de cuatro siglos de duración, que ha sido recientemente presentado a la atención de los teóricos de la evolución por David Sloan Wilson y Elliot Sober:

> Los hutteritas son una secta religiosa fundamentalista que se originó en Europa en el siglo XIX y que emigró a Norteamérica en el siglo XVI, para escapar de la llamada a filas. Los hutteritas se consideran a sí mismos como el equivalente humano de una colmena. Practican la comunidad de bienes (no hay propiedad privada) y también cultivan una actitud psicológica de extrema despersonalización… El nepotismo y la reciprocidad, los dos principios que la mayor parte de los evolucionistas utilizan para explicar la conducta prosocial en los humanos, son menospreciados por los hutteritas por inmorales. Dar a los demás debe hacerse sin tener en cuenta el parentesco y sin ninguna expectativa de retorno (Wilson y Sober 1994:602.).

Al contrario que muchas sectas, los hutteritas han tenido mucho éxito en la propagación de sus grupos a lo largo de los siglos incrementando su extensión territorial y su población global, según Wilson y Sober: «En el Canadá actual los hutteritas habitan en granjas marginales sin el beneficio de la moderna tecnología y casi puede asegurarse que desplazarían a la población no hutterita en ausencia de leyes que restrinjan su expansión» (Wilson y Sober 1994:605).

Los hutteritas pueden tener más de cuatro siglos de antigüedad pero este no es tiempo suficiente en el calendario genético, así que no es probable que *alguna* de las diferencias significativas entre sus grupos y los grupos a los que pertenecemos el resto sean genéticamente transmitidas (intercambiar niños hutteritas por otros probablemente no interferiría de modo significativo en la adaptación del grupo en las colonias hutteritas. Los hutteritas simplemente sacan partido, gracias a una herencia de transmisión cultural, de tendencias que son compartidas por el género humano). De este modo, los hutteritas son un ejemplo de cómo la evolución cultural puede crear un nuevo grupo y, lo que es particularmente atractivo, desde un punto de vista evolucionista, es su método de escisión:

Como una colmena de abejas, las hermandades hutteritas se dividen cuando alcanzan un gran tamaño; una mitad permanece en el sitio original y la otra se traslada a un nuevo lugar que ha sido preseleccionado y acondicionado. En la preparación para la escisión, la colonia se divide en dos grupos que son iguales respecto al número, edad, sexo, habilidades y compatibilidad personal. La colonia entera empaqueta sus pertenencias y una de las listas es extraída por lotería el día de la separación. Difícilmente puede ser más completa la similitud del procedimiento con las reglas de la meiosis (Wilson y Sober 1994:604).

¡El velo darwiniano de la ignorancia en acción! Pero esto no es suficiente, por sí mismo, para asegurar la solidaridad del grupo, dado que los seres humanos, incluso aquellos que han vivido toda su vida en una comunidad hutterita, no son sistemas intencionales balísticos, sino sistemas intencionales guiados, y la guía ha de ser facilitada diariamente. Wilson y Sober citan a Ehrenpreis, uno de los primeros líderes de la secta: «Una y otra vez observamos que el hombre con su actual naturaleza encuentra muy difícil practicar una vida de verdadera comunidad». Continúan aportando más citas en las cuales Ehrenpreis subraya cómo han de ser de explícitas y enérgicas las prácticas de los hutteritas para contrarrestar esta tendencia demasiado humana. Estas declaraciones dejan claro que, por una vía o por otra, la organización social hutterita es el resultado de prácticas culturales vigorosamente dispuestas en orden de batalla *contra* las verdaderas características de la naturaleza humana que Wilson y Sober desean negar o minimizar: egoísmo e imparcialidad para razonar. Si el pensamiento grupal perteneciera realmente a la naturaleza humana como a Wilson y Sober les gustaría creer, los padres hutteritas y los ancianos no tendrían nada que decir. (Comparemos esto con el caso en el cual hay verdaderamente una predisposición genética en nuestra especie: ¿con qué frecuencia han oído de padres que engatusen a sus hijos para comer más dulces?).

Wilson y Sober están en lo cierto al presentar los ideales hutteritas como

la esencia de una organización organísmica, pero la gran diferencia es que para la gente —al contrario de las células en nuestros cuerpos o las abejas en su colmena— siempre existe la opción de rehusar. Y esto, a mi parecer, es la última cosa que deseamos destruir en nuestra ingeniería social. Los hutteritas no están de acuerdo, aparentemente, al igual que, por lo que veo, los portadores de muchos memes no occidentales[163].

¿Le *gusta* a usted la idea de convertirnos, nosotros y nuestros hijos, en esclavos por el *summum bonum* de nuestros grupos? Esta es la dirección en la cual se han movido siempre los hutteritas y, según cuentan Wilson y Sober, consiguen un éxito impresionante, aunque sólo sea a costa de prohibir el libre intercambio de ideas y desalentar pensar por uno mismo (lo que ha de distinguirse de ser egoísta). Cualquier obstinado librepensador será conducido ante la congregación y amonestado con firmeza; «si persiste en su obstinación y rehúsa escuchar incluso a la Iglesia, entonces hay una sola respuesta a esta situación, y es romper con él y excluirlo». Un régimen totalitario (incluso un totalitarismo grupal) es extremadamente vulnerable a la disuasión, casi exactamente del mismo modo que un grupo altruista es vulnerable a los oportunistas. Esto no significa que la razón esté siempre del lado de la defección. No es así. La razón está siempre del lado en el que se mantienen abiertas las opciones y se *revisan* los diseños. Esto es generalmente, pero no siempre, un importante hecho que había sido señalado por el economista Thomas Schelling [1960], el filósofo Derek Parfit [1984] y otros, en sus discusiones de las condiciones bajo las cuales sería racional para un agente racional convertirse a sí mismo (temporalmente) en irracional. (Por ejemplo, si uno desea convertirse en un objetivo desfavorable para una extorsión, debe convencer, de algún modo, al mundo de que es impenetrable a la razón, por lo que el mundo no intentará hacerle ofertas que no pueda razonablemente rehusar).

Hay circunstancias —circunstancias extremas, como hacen notar Wilson y Sober— en las que podemos razonablemente reducir la libertad de pensamiento, aunque los hutteritas se oponen a esta libertad: desaconsejan leer cualquier libro y escuchar a quien se quiera. La conservación de un estado prístino —como el de los hutteritas— sólo puede conseguirse con el cuidadoso control de los canales de comunicación. Esta es la razón por la que la solución organísmica no es una solución para los problemas de la sociedad humana. Los hutteritas son un curioso ejemplo de reduccionismo insaciable, no porque ellos individualmente sean insaciables —aparentemente se comportan de manera opuesta a la insaciabilidad—, sino porque su solución

al problema de la ética es una solución drásticamente hipersimplificada. Ellos son, sin embargo, muy buen ejemplo del poder de los memes para infectar un grupo de comunicadores mutuos, de tal modo que todo el grupo dirige sus esfuerzos a asegurar la proliferación de *aquellos memes* al coste que sea entre ellos mismos**[164]**.

En la próxima sección, estudiaremos más de cerca lo que la sociobiología es y no es, lo que podría y no podría ser, pero antes de abandonar el tema del reduccionismo ético insaciable, debemos detenernos para considerar una antigua especie de este desagradable meme con muchas subvariedades: la religión. Si queremos presentar un claro ejemplo de la falacia naturalista, difícilmente podríamos encontrar uno mejor que ponga al descubierto la práctica de intentar la justificación de un precepto ético, un «debe», citándolo como un «es»: así lo dice la Biblia. Ante esta postura, como a Wilson y Sober, debemos decir: ¿Y qué? ¿Por qué debe suponerse que los hechos relatados en la Biblia —si realmente se trata de hechos— (o en otro texto sagrado, añado rápidamente) facilitan una justificación más satisfactoria para los principios de la ética que los hechos citados por Darwin en *El origen de las especies*? Ahora bien, si se cree que la Biblia (o cualquier otro texto sagrado) es *literalmente* la palabra de Dios, y que los seres humanos han sido puestos en la Tierra por Dios para cumplir su mandato, convirtiendo la Biblia es una especie de manual para el uso de las herramientas de Dios, entonces se está en posesión de los fundamentos para creer que los principios éticos hallados en la Biblia tienen una garantía especial con la que no cuentan otros escritos. Si, por otro lado, se cree que la Biblia, como la *Odisea* de Homero, *El Paraíso Perdido* de Milton y *Moby Dick* de Melville, es realmente un producto de la cultura humana, nada milagroso, surgido de uno o más autores humanos, entonces no se le otorgará autoridad a la Biblia, más allá de la tradición y de cualquiera de los argumentos generados por su propia coherencia. Este es, debería ser obvio, el indisputable punto de vista de los filósofos que se ocupan hoy de la ética, tan incontrovertible que, si intentamos refutar una afirmación en la literatura ética contemporánea señalando que la Biblia dice otra cosa, nos encontraremos con sorprendidas miradas de incredulidad. «¡He aquí la falacia naturalista!», pueden decir los filósofos de la ética. «¡De este tipo de “es” no puede derivar “debe”!». (Así que no debemos esperar que los filósofos acudan a defendernos si afirmáramos que la religión es una fuente de sabiduría ética superior en cualquier sentido a la ciencia).

¿Significa esto que los textos religiosos no tienen valor como guías para la ética? Desde luego que no. Estos textos son magníficas fuentes de perspicaz penetración en la naturaleza humana y en las posibilidades de los códigos éticos. Así como no debe sorprendernos descubrir que la antigua medicina popular tiene mucho que enseñar a la moderna medicina con su alta tecnología, no debe sorprendernos que estos grandes textos religiosos contengan versiones de los mejores sistemas éticos que la cultura humana ha diseñado. Pero, como sucede con la medicina popular, debemos examinar estos textos cuidadosamente y no aceptar nada por fe. (¿Cree alguien que es muestra de sabiduría comer hongos «santos», sencillamente porque tradiciones milenarias sostienen que nos ayudan a ver el futuro?). El punto de vista que estoy exponiendo corresponde al denominado, con frecuencia, «humanismo secular». Si uno tiene al humanismo secular como su «fantasma», no debe concentrar toda su energía en atacar a los sociobiólogos o conductistas o filósofos académicos, porque ellos son únicamente una fracción de un porcentaje de pensadores influyentes que firme y tranquilamente creen que la ética no debe *asentarse* —como mucho guiada— en doctrinas religiosas. Esta es, naturalmente, la opinión reinante en el Congreso de los Estados Unidos y en los tribunales; citar la Constitución tiene más importancia que citar la Biblia, y así debe ser.

Los culpables de la mala fama del humanismo secular son, a menudo, los que se titulan humanistas seculares, reduccionistas insaciables, de un tipo o de otro, impacientes con las complejidades de las tradiciones antiguas, irrespetuosos con las genuinas maravillas que cabe saborear en la rica tradición cultural de otros. Si piensan estos humanistas que todas las cuestiones éticas pueden reducirse a una definición o a unas pocas definiciones simples (si esto es malo para el medio ambiente, es malo; si esto es malo para el arte, es malo; si esto es malo para los negocios, es malo), en este caso no son mejores filósofos de la ética que Herbert Spencer y los darwinistas sociales. Pero si hacemos la contrapropuesta, completamente apropiada, de que la vida es más complicada, hemos de ser cuidadosos para no convertirla en una actitud obstructiva ante la investigación, en lugar de en una llamada a favor de una investigación del problema de la ética más cuidadosa. De otra manera, nos volveremos a situar de nuevo en un movimiento pendular, en el lado de la desesperanza.

¿A qué se parecería, entonces, una investigación más cuidadosa? La tarea a la que nos enfrentamos es, todavía, la tarea a la que se enfrentaron Hobbes y Nietzsche: *de alguna manera* tenemos que haber evolucionado hasta

convertirnos en seres que pueden tener una consciencia que, como dice Nietzsche (*Más allá del bien y del mal*, §98), nos besa mientras nos muerde. Un forma atractiva de plantear la cuestión es imaginar que somos un selector *artificial* de gente altruista. Como un criador de ganado doméstico, palomas, o perros, podemos observar de cerca a nuestro grupo, anotando en un libro mayor los que son díscolos y los que se comportan amablemente y, mediante intromisiones de distinto tipo, disponer las cosas de modo que los que se comportan bien tengan más descendencia. A su debido tiempo, debemos ser capaces de hacer evolucionar el grupo hacia una población de gente buena, en el supuesto de que una tendencia a esta bondad pudiera ser representada en un genoma. No cabe pensar que lo que hacemos es una selección por un «módulo ético» que es diseñado precisamente *para* dar respuestas correctas a cuestiones éticas. Cualquier modelo o artefacto puede tener, de manera singular o en coalición, el efecto (o producto colateral o bonificación) de favorecer las elecciones altruistas en el momento de tomar la decisión. Después de todo, la lealtad de los perros a los seres humanos es, según parece, el resultado de una selección inconsciente realizada por nuestros antepasados. Es concebible que Dios pudo haber hecho esto para nosotros, pero supongamos que nuestro deseo es eliminar el Intermediario y explicar la evolución de la ética por la selección *natural*, no la selección *artificial*. ¿Puede haber algunas fuerzas ciegas, carentes de visión del futuro, algún grupo de circunstancias naturales, que puedan conseguir lo mismo?

Todo el mundo puede ver que hay caminos tortuosos y graduales por los cuales hemos penetrado, por nuestros propios medios, en el territorio de la moral genuina, mediante una serie de muy pequeños cambios, y no de un solo golpe. Podemos comenzar con la «inversión de los padres» (Trivers 1972). Es indiscutible que las mutaciones que generan criaturas que invierten más energía y tiempo en cuidar a sus hijos, pueden evolucionar, en la mayoría de las circunstancias. (Recuérdese que sólo algunas especies se implican en la inversión de los padres. Esta inversión de tiempo y energía no es una opción para especies en las que la criatura sale del cascarón después que los padres han muerto, y la razón de esta pauta de comportamiento tan fundamentalmente diferente de los padres ha sido bien investigada)**[165]**. Ahora bien, una vez que la inversión de los padres en su propia cría está asegurada para una especie, ¿cómo extender el círculo (Singer 1981)? Gracias al trabajo pionero de Hamilton (1964) sobre la «selección de la descendencia» y la «capacidad de sobrevivir del individuo que *incluye* a descendientes», parece también indiscutible que las mismas condiciones que favorecen el

sacrificio por una cría, también favorecen, en un grado matemáticamente preciso, los sacrificios por parientes más lejanos: crías que ayudan a los padres, gemelos que se ayudan mutuamente, tías que ayudan a sobrinos, y así sucesivamente. Pero, una vez más, es importante recordar que las condiciones bajo las cuales esa ayuda es reforzada por la evolución no sólo no son universales, sino relativamente raras.

Como George Williams [1988] ha señalado, no sólo es común el canibalismo (comerse los coespecíficos, incluso parientes cercanos), sino que en muchas especies matar a los gemelos o «gemelocidio» (no lo podemos llamar crimen dado que ellos no saben lo que hacen) es casi la regla, no la excepción. (Por ejemplo, cuando nacen en el mismo nido dos o más polluelos de águila, el primero que rompe el cascarón es muy probable que mate a su gemelo más joven, si le es posible, lanzando el huevo fuera del nido o incluso empujando fuera al polluelo que acaba de romper el cascarón). Cuando un león adquiere una nueva leona que está aún criando cachorros nacidos de un apareamiento anterior, la primera decisión del nuevo trato entre ambos animales es matar a estos cachorros, para que la leona puede entrar más rápidamente en un período fértil. Los chimpancés son conocidos por implicarse en combates a muerte contra otros chimpancés, y los langures machos, monos de la India, matan a menudo a los hijos de otros machos para conseguir acceso reproductivo a las hembras (Hrdy 1977). De modo que, incluso nuestros parientes más cercanos, se implican en conductas horribles. Williams llama la atención sobre el hecho de que en todas las especies de mamíferos que han sido cuidadosamente estudiadas, la tasa de participación de sus miembros en matanzas de coespecíficos es varios miles de veces mayor que la más alta tasa de homicidios en cualquier ciudad norteamericana[166].

Este sombrío mensaje acerca de nuestros peludos amigos es contrarrestado a menudo, y las presentaciones populares de la naturaleza (documentales de televisión, artículos en revistas y libros de divulgación) practican con frecuencia la autocensura para evitar impresionar a los remilgados. Hobbes estaba en lo cierto: la vida en el estado de naturaleza es desagradable, brutal y corta, para prácticamente todas las especies no humanas. Si «hacer lo que es natural» significa lo que virtualmente hacen las otras especies animales, sería peligroso para la salud y el bienestar de todos nosotros. Es célebre la frase de Einstein: «el querido Dios es sutil pero no malicioso»; Williams volvió del revés esta observación: la madre naturaleza

es desalmada —incluso perversa—, aunque es extremadamente estúpida. Y como ocurre a menudo, Nietzsche centra la cuestión y da su toque especial:

¿Debo vivir de «acuerdo con la naturaleza»? ¡Oh, vosotros nobles estoicos, qué palabras tan decepcionantes! Imagina un ser como la naturaleza, derrochador desmedido, indiferente más allá de toda medida, sin propósitos ni consideración, sin piedad ni justicia, al mismo tiempo fértil y desolado e incierto; imagina la indiferencia en sí misma como un poder; ¡cómo podéis vivir de acuerdo con esta indiferencia! (*Más allá del bien y del mal*, §9).

Más allá de la capacidad de sobrevivir del individuo que *incluye* a descendientes, se encuentra el «altruismo recíproco» (Trivers 1971) en el cual organismos con parentesco lejano o sin parentesco —no necesitan ni siquiera ser de la misma especie— pueden llegar a unos acuerdos mutuamente beneficiosos de *quid pro quo*, el primer paso hacia un humano mantenimiento de promesa. Habitualmente se objeta que la expresión «altruismo recíproco» es inapropiada, dado que no se trata *realmente* de altruismo sino de un ilustrado interés propio, de una forma o de otra: literalmente, si rascas mi espalda, yo te rascaré la tuya, en el caso de acuerdos para el aseo que son un ejemplo simple y socorrido. Esta «objeción» no tiene en cuenta que debemos acercarnos a lo que verdaderamente nos interesa dando pequeños pasos, y el «altruismo recíproco», innoble (o justamente a-noble) quizás es un útil escalón en la progresión. Requiere habilidades cognitivas avanzadas, una memoria bastante específica capaz de reidentificar deudores y acreedores, y la capacidad de descubrir un tramposo, por ejemplo.

Ir más allá de las formas más interesadas económicamente y más brutales del altruismo recíproco, hacia un mundo en el cual la genuina confianza y el sacrificio son posibles, es una tarea que ha comenzado a ser explorada teóricamente. El primer gran paso fueron los torneos con el «Dilema del Prisionero» de Robert Axelrod (Axelrod y Hamilton 1981, Axelrod 1984), los cuales invitaron a todos los interesados a enviar estrategias —algoritmos— para competir entre todos los participantes en un reiterado Dilema del Prisionero. (Entre las muchas discusiones de este dilema, dos de las mejores son las de Dawkins 1989a:cap. 12, y Poundstone 1992). La estrategia del ganador ha llegado a ser famosa: golpe por golpe, la cual simplemente copia el movimiento previo del «oponente», cooperando como recompensa a su pasada cooperación y traicionándolo como represalia contra cualquier traición. El básico golpe por golpe se presenta en una variedad de subespecies. En el amable golpe por golpe uno comienza cooperando y entonces le hace al otro lo que el otro le hace a uno en el movimiento previo. Como puede verse fácilmente dos amables jugadores de golpe por golpe,

jugando uno como contrincante del otro lo hacen espléndidamente, cooperando de manera indefinida, pero un amable jugador de golpe por golpe que se encuentre frente a un malintencionado jugador de golpe por golpe que introduce una no provocada defección en cualquier punto del juego inicia, como represalia, una debilitadora serie de defecciones sin fin (esto sirve para ambos, naturalmente, mientras no se olviden de los pasos que cada uno de ellos van dando).

Las situaciones simples exploradas por el torneo inicial de Axelrod han abierto el camino a situaciones más complejas y realistas. Nowak y Sigmund [1993] han encontrado una estrategia que supera al golpe por golpe en una importante variedad de circunstancias. Kitcher [1993] examina un mundo de juegos del Dilema del Prisionero *no coercitivos* (si no le gusta un concreto oponente, decide no jugar). Kitcher demuestra, con datos matemáticos muy detallados, cómo los «altruistas discriminantes» (que guardan una lista de a quienes derrotaron en el pasado) pueden florecer bajo ciertas —no todas— condiciones, y también comienza a clasificar las condiciones bajo las cuales, variando las políticas de perdón y de olvido, puede mantener su propiedad contra la siempre presente perspectiva de un resurgimiento de tipos antisociales. Es particularmente fascinante, en la dirección abierta por el análisis de Kitcher, la emergencia de grupos en los cuales el fuerte y el débil tenderían a segregarse ellos mismos y preferirían cooperar con los de su propia clase.

¿Podría esto preparar el escenario para algo así como el intercambio de los valores de Nietzsche? Han sucedido cosas extrañas. Stephen White (no publicado) ha comenzado a investigar las importantes complejidades del Dilema del Prisionero jugado entre más personas. (Este es el juego que conduce a la tragedia de las propiedades comunales, que da lugar, simultáneamente, al agotamiento de las reservas de pesca en nuestros océanos y de los bosques de árboles altos). Como ha señalado Kitcher, los escenarios simples son tratables analíticamente —las ecuaciones de interacción y sus esperados productos pueden resolverse por cálculo matemático— pero cuando añadimos más realismo y, en consecuencia, complejidad, la *solución* directa de las ecuaciones llega a ser *no factible*, de modo que tendremos que recurrir a los métodos indirectos de la simulación en el ordenador. En esta simulación, se reúnen cientos o miles de individuos imaginarios provistos con docenas o cientos o miles de estrategias u otras propiedades y deja que el ordenador los sustituya jugando miles o millones de juegos entre sí, guardando los resultados**[167]**.

Ésta es una rama de la sociobiología o ética evolucionista que nadie debe ridiculizar. Directamente somete a prueba las corazonadas, tales como aquellas de Hobbes y Nietzsche que son vías naturales, reforzables evolutivamente, hasta donde nos encontramos hoy. Podemos estar absolutamente seguros de que esto es verdad, desde donde estamos, pero lo que esta investigación promete clarificar es cuánto trabajo de I+D, y de qué tipos, fue necesario para traernos hasta aquí. En un extremo, podría resultar que hay un impresionante cuello de botella; una completamente improbable pero crucial serie de afortunados acontecimientos que fueron necesarios. (El análisis de White ofrece algunas plausibles razones para creer que las condiciones eran realmente muy estrictas). En el otro extremo, podría resultar que hay una más que amplia cuenca de atracción que llevará a casi todas las criaturas cognitivamente complejas, cualesquiera que sean sus circunstancias, a integrarse en sociedades con códigos éticos reconocibles. Será fascinante ver lo que las simulaciones en el ordenador a gran escala de estas complejas relaciones intersociales nos dicen acerca de las limitaciones en la evolución de la ética. Pero podemos estar ya virtualmente seguros de que el reconocimiento mutuo y la capacidad de comunicar una promesa —subrayadas ambas tanto por Hobbes como por Nietzsche— son condiciones necesarias para la evolución de la moral. Es concebible, aunque improbable con la evidencia actual, que los ballenas y los delfines, o los grandes simios, cumplan estas necesarias condiciones, pero ninguna de las otras especies se acercan tanto a la exhibición de esos tipos de cognición social de los que depende la verdadera moral. (Mi corazonada pesimista es que la principal razón para que todavía no hayamos descartado a los delfines y a las ballenas como seres morales de las profundidades es que son muy difíciles de estudiar en su vida natural. La mayoría de la evidencia acerca de los chimpancés —algunas de estas evidencias han sido autocensuradas por los investigadores durante años— indica que son verdaderos ciudadanos del estado de la naturaleza de Hobbes, mucho más desagradable y brutal de lo que a muchos de nosotros nos gustaría creer).

4. *Sociobiología: lo bueno y lo malo, el bien y el mal*

> El cerebro humano funciona como funciona. Desear que funcione de algún modo como un atajo para justificar algún principio ético, socava tanto a la ciencia como a la ética (¿qué sería de tal principio si los hechos demostraran lo contrario?).
>
> Steven Pinker,

> La sociobiología tiene dos caras. Una observa la conducta social de los animales no humanos. Los ojos cuidadosamente enfocados y los labios fruncidos juiciosamente. Las declaraciones se hacen con cautela. La otra cara está casi oculta tras un megáfono. Las declaraciones acerca de la naturaleza humana se proclaman ruidosamente y con gran excitación.
>
> Philip Kitcher,
> *Vaulting Ambition*

Otra parte de nuestra investigación sobre la naturaleza humana, como una base naturalista que sirva de fundamento a un pensamiento ético, debería comenzar con el hecho indiscutible de que nosotros, seres humanos, somos productos de la evolución, y considerar las limitaciones con las que nacemos y las variaciones existentes entre nosotros que puedan poseer relevancia ética. Mucha gente cree, aparentemente, que la ética se encontraría en graves dificultades si resultase que nosotros, seres humanos, no estamos, como dice la Biblia, tan sólo un poco por debajo de los ángeles. Si no somos todos perfecta e igualmente racionales, y perfecta e igualmente maleables por la educación, e igualmente capaces en todos los otros aspectos, entonces nuestras premisas fundamentales de igualdad y perfectibilidad se encuentran en peligro. Si esto fuera verdad, sería demasiado tarde para salvarnos, porque ya conocemos suficientemente las fragilidades y las diferencias humanas como para sostener esta visión. Pero hay modos más razonables de ver la cuestión que también se encuentran en peligro debido a los descubrimientos de los científicos (no precisamente evolucionistas).

No hay duda de que el conjunto de hechos que podemos aprender acerca de un individuo, o de un grupo de individuos (mujeres, gente de ascendencia asiática, etc.) puede afectar profundamente a nuestra tendencia a considerarlos y a tratarlos. Si me entero de que Sam es esquizofrénico o profundamente retrasado, o sufre de mareos y desmayos periódicos, no lo contrataré como conductor del autobús escolar. Cuando pasamos de hechos específicos de individuos a generalizaciones acerca de grupos de individuos, la situación es más complicada. ¿Cuál es la respuesta, justa y razonable, de las compañías de seguros ante los datos actuariales acerca de las diferentes expectativas de vida de hombres y mujeres? ¿Es honesto ajustar las primas de acuerdo con estas expectativas? O ¿deben tratarse ambos sexos como en el departamento de primas y aceptar como honesta la tasa diferencial de beneficios entre hombres y mujeres? Con respecto a las diferencias voluntariamente adquiridas

(fumadores frente a no fumadores, por ejemplo), ¿consideramos honesto que los fumadores paguen por su hábito primas mayores?, pero ¿qué pasa con las diferencias respecto a los distintos lugares de nacimiento? Los afroamericanos poseen generalmente, como grupo, una tendencia mayor a la presión arterial elevada, la diabetes es más frecuente que el término medio entre los hispanos, y los blancos son más propensos al cáncer de piel y a la fibrosis quística (Diamond 1991). ¿Deben reflejarse estas diferencias en el cálculo de su prima para el seguro de enfermedad? Aquellos niños cuyos *padres* fuman en casa mientras se encuentran en período de crecimiento, corren un riesgo más elevado de padecer trastornos respiratorios aunque la culpa no sea de ellos. Los jóvenes, como grupo, son conductores menos seguros que las jóvenes. ¿Cuáles de estos hechos deben ser tenidos en cuenta, hasta qué punto y por qué? Aun cuando nos ocupemos de hechos referentes a individuos concretos, más que a tendencias estadísticas, hay dilemas en abundancia: ¿están autorizados los patronos —u otras personas— para conocer si he estado casado, si tengo antecedentes penales, historia de conductor seguro de automóviles o si practico el submarinismo? ¿Hay diferencia, en principio, entre hacer públicas las calificaciones obtenidas en la escuela por una persona y dar a conocer información acerca del coeficiente intelectual de la misma persona?

Estos son difíciles problemas éticos. La ciudadanía debate actualmente varias restricciones sobre qué información sobre los individuos pueden investigar los patronos, el gobierno, las escuelas, las compañías de seguros y otros organismos e instituciones y hay un corto paso para llegar a la conclusión de que lo mejor sería que la ciencia no investigara cierto tipo de información. Acaso sería mejor que nos mantuviéramos ignorantes acerca de cosas tales como si existen grandes diferencias entre los cerebros de los hombres y los de las mujeres, si es un gen el que predispone a la dislexia, a la violencia, al genio musical o a la homosexualidad. No se debe rechazar a la ligera esta sugerencia. Si alguna vez nos hemos preguntado si hay hechos acerca de nosotros mismos (de nuestra salud, competencia, perspectivas) que sería mejor no conocer y decidimos que así sea, debemos estar preparados para considerar seriamente la sugerencia de que la mejor, y quizá la única vía, para asegurar que tales hechos no sean impuestos a la gente, sería prohibir las investigaciones que los pudieran descubrir[168].

Por otra parte, si no investigamos estos temas perderemos importantes oportunidades. La sociedad tiene un gran interés en mantener archivos de los arrestos por conducción en estado de ebriedad de potenciales conductores de

autobuses escolares y darlos a conocer a quienes tienen que tomar las decisiones, y tiene el mismo gran interés en descubrir cualquier tipo de hechos entre sus miembros que puedan potenciar nuestras vidas o proteger a toda la sociedad de miembros concretos de ésta. Esto es lo que hace que las decisiones de investigar que tomemos sean tan cruciales y tan propicias a generar controversias. No es sorprendente que la investigación sociobiológica se lleve a cabo en una atmósfera de incesante preocupación con tendencia a convertirse en alarma, y cuando se produce la escalada, como sucede a menudo, la propaganda entierra la verdad.

Comencemos con el término «sociobiología». Cuando E. O. Wilson lo acuñó, intentaba cubrir el espectro total de las investigaciones biológicas concernientes a la evolución de las interrelaciones entre organismos en parejas, grupos, manadas, colonias, naciones. Los sociobiólogos estudian las relaciones entre las termitas en un montículo, las crías del cuco y sus engañados padres adoptivos, los miembros de los grupos matriarcales de elefantes, las bandas de monos, los elefantes marinos machos y sus harenes, así como las parejas humanas, familias, tribus y naciones. Pero, como dice Kitcher, la sociobiología de los animales no humanos siempre se ha llevado a cabo con mayor cuidado y precaución. (Véase también Ruse 1985). De hecho, se incluyen en este terreno algunos de los más importantes y ampliamente conocidos avances en la reciente biología teórica, tales como los clásicos artículos de Hamilton, Trivers y Maynard Smith.

Podría decirse que fue Hamilton el que inauguró este campo del conocimiento con su introducción del concepto de la selección de la parentela, el cual solucionó, entre otras cosas, muchos de los rompecabezas darwinianos acerca de la *eusocialidad* en los insectos: el modo en que hormigas, abejas y termitas viven «sin egoísmo» en grandes colonias, la mayor parte de los individuos como sirvientes estériles de una singular reina fértil. Pero la teoría de Hamilton no resuelve todos los problemas y entre las importantes contribuciones de Richard Alexander se encuentra su descripción de las condiciones bajo las cuales los *mamíferos* eusociales pueden evolucionar; una «predicción» asombrosamente confirmada por los consiguientes estudios de la maravillosa rata-topo lampiña de Sudáfrica (Sherman, Jarvis y Alexander 1991). Este fue un impresionante triunfo del razonamiento adaptacionista que merece ser más ampliamente conocido. Como Karl Sigmund señaló, las ideas de Hamilton

condujeron a un notable descubrimiento cuando en 1976 el biólogo americano R. D. Alexander disertó sobre las castas estériles. Es bien conocido que estas castas existen en las hormigas, abejas y termitas,

pero no en clase alguna de vertebrados. Alexander, en una variante de experimento mental, jugó con la noción de un mamífero que fuera capaz de evolucionar hacia una casta estéril. Sería necesario, como en las termitas, un nido expandible que permitiera almacenar una gran cantidad de alimentos y que, al mismo tiempo, fuera un refugio contra los predadores. Por razones de tamaño, una localización bajo la corteza (como la de los presuntos insectos ancestros de las termitas) no sería apropiada. Por el contrario, unas galerías subterráneas repletas de grandes tubérculos cumplirían el papel a la perfección. El clima sería tropical; el suelo (¡más que una alusión a Sherlock Holmes en este punto!) arcilloso. Un ingenioso ejercicio de ecología de gabinete. Pero, después de su conferencia, se le hizo saber a Alexander que este hipotético animal vivía en África: era la rata-topo lampiña, un pequeño roedor estudiado por Jennifer Jarvis (Sigmund 1993:117).

La rata-topo lampiña es tremendamente fea y extraña, un experimento mental de la madre naturaleza que rivaliza con cualquier fantasía filosófica. Estos animales son genuinamente eusociales. La reina de la rata-ropo lampiña es la única hembra que cría y mantiene bajo sus órdenes al resto de la colonia, mediante la liberación de feromonas que suprimen la maduración de los órganos reproductores de las restantes hembras. La rata-topo lampiña es coprófaga, ya que con regularidad ingieren sus propias heces, y cuando la rata reina preñada, grotescamente hinchada, no puede alcanzar su propio ano, ingiere las heces de sus asistentes. (¿Tienen bastante con esto? Hay mucho más, bastante más, todo muy recomendable para aquellos cuya curiosidad excede su escuprulosidad). Un regalo ha sido lo aprendido con las ratas-topos lampiñas y otras especies no humanas, utilizando las técnicas darwinianas de la ingeniería revertida —en otras palabras, usando el adaptacionismo— y seguramente se ha de aprender mucho más. El importante trabajo del propio E. O. Wilson (1971) sobre los insectos sociales es merecidamemte famoso en el mundo entero, y hay hoy, literalmente, cientos de otros magníficos sociobiólogos de animales. (Véanse, por ejemplo, las clásicas antologías de Clutton-Brock y Harvey 1978, Barlow y Silverberg 1980 y del King’s College Sociobiology Group 1982). Desafortunadamente, todos ellos trabajan envueltos en una nube de sospecha, formada por la escalada de unos cuantos sociobiólogos humanos con sus insaciables afirmaciones (a través de sus megáfonos, como sugiere Kitcher), las cuales resuenan con el eco de sus oponentes, en otra escalada de condenas generalizadas. Esto es, sin duda, una consecuencia desafortunada, pues, como en cualquier otra área legítima de la ciencia, parte de estos trabajos es muy importante, parte es buena, parte es buena pero falsa, y parte es mala, aunque ninguna representa el mal. Que los que estudian con seriedad los sistemas de apareamiento, el despliegue de los cortejos previos, la territorialidad y cosas parecidas, en las especies no humanas, deban ser cortados por el mismo patrón que los más flagrantes

transgresores en la sociobiología humana es, al mismo tiempo, un extravío de la justicia y una grave tergiversación de la ciencia.

Pero, ninguna de las dos «partes» ha cumplido con su deber. Desafortunadamente, una mentalidad de asedio ha hecho que los mejores sociobiólogos sean algo reticentes a criticar la mala calidad de los trabajos de algunos de sus colegas. Aunque en Maynard Smith, Williams, Hamilton y Dawkins se pueden leer, con cierra frecuencia, correcciones de algunos inocentes errores en argumentos y puntúalizaciones acerca de sus complicaciones —en pocas palabras, haciendo las correciones que son habituales en las ponencias científicas—, todos ellos se han abstenido ante los más insignes pecados incluidos en los trabajos de aquellos que, entusiásticamente, hacen mal uso del propio buen trabajo de los citados. Sin embargo, Donald Symons [1992] es una gratificante excepción, y hay otras. Quiero señalar una *importante* fuente de razonamiento incorrecto, muy extendido en la sociología humana, y que raramente es utilizado por los mismos sociobiólogos, quizá porque Stephen Jay Gould lo ha criticado y para ellos sería odioso tener que conceder que Gould está en lo cierto acerca de alguna cosa. Gould lo está sin duda en este punto, como lo está Philip Kitcher [1985b], quien desarrolló esta crítica con mucho más detalle. He aquí la versión de Gould, que es algo más difícil de entender. (Al principio, no me pareció que pudiera leerse con simpatía, y hube de preguntar a Ronalf Admunsen, un excelente filósofo de la biología, que me explicara lo que Gould quería decir. ¡Tuvo éxito!).

El fundamento estándar de las historias del darwinismo no se aplica a los humanos. Este fundamento es la implicación: si es adaptativa, entonces es genética, pues la inferencia de la adaptación es generalmente la única base de una historia genética y el darwinismo es una teoría de cambios genéticos y variaciones en la población (Gould 1980c:259).

¿Qué significa esto? Gould no dice, como podría parecer a primera vista, que la inferencia adaptacionista no se aplique a los humanos. Dice que, dado que en los humanos (y sólo humanos) hay siempre *otra* posible fuente de la adaptación en cuestión denominada cultura, no se puede tan fácilmente inferir que haya habido una evolución genética para la característica concreta que se considera. Incluso en los animales no humanos, la inferencia de la adaptación con bases genéticas es arriesgada, cuando la adaptación considerada no es una característica anatómica, sino un patrón de conducta que es, obviamente, una buena maña. Para esto existe otra posible explicación: la especie, en general, *no es estúpida*. Como hemos visto con frecuencia, mientras más obvio es el movimiento, menos segura será la inferencia de que esta característica ha de

haber sido copiada de los predecesores, y transmitida específicamente por los genes.

Hace muchos años, jugaba mi primer videojuego en el ordenador del laboratorio de inteligencia artificial en el MIT: se llamaba «Maze War» ('Guerra en el laberinto’) y más de una persona podía jugar al mismo tiempo, cada una situada ante un terminal independiente, conectado a un ordenador de tiempo compartido. En la pantalla se podía ver dibujado en perspectiva, con un simple trazo, un laberinto, en el cual yo, el observador, estaba situado. Podía ver ante mí pasillos que se abrían a derecha o a izquierda, y presionando teclas me podía mover hacia delante o hacia atrás o girar noventa grados a la izquierda o a la derecha. Otra tecla correspondía al gatillo del arma que disparaba hacia delante. Los otros jugadores se encontraban ante el mismo laberinto virtual, deambulando por allí, buscando a alguien a quien disparar y con la esperanza de no ser alcanzados por un disparo. Si uno de los jugadores se cruzaba en mi camino, aparecía como una simple figura de cartón, a la que yo esperaba alcanzar con un disparo antes de que se volviera y me disparara. Después de unos minutos de frenético juego, durante el cual fui «herido» por la espalda varias veces, mi creciente paranoia era tan inconfortable que busqué alivio: hallé un *cul-de-sac* en el laberinto y, allí encerrado, me sentí relativamente tranquilo con mi dedo en el gatillo. Me sorprendió entonces que hubiera adoptado la conducta de la morena, que espera pacientemente en su bien protegido agujero a que pase por delante algo que valga la pena para salir nadando en su persecución.

Ahora bien, mi comportamiento en esta ocasión, ¿nos proporciona alguna razón para suponer que existe en el *Homo sapiens* una predisposición genética a adoptar la conducta de la morena? ¿Fue el estrés de la ocasión el que hizo aflorar algún antiguo comportamiento que permanecía durmiendo en mis genes, desde los días en los que mis ancestros eran aún peces? Desde luego que no. La estrategia es demasiado obvia. Aparece como un movimiento forzado, aunque sea, al menos, una buena maña. No nos sorprendería encontrar a marcianos que corrieran a esconderse en la seguridad de sus cuevas en Marte, y las posibilidades de que, debido a este descubrimiento nuestro, los marcianos tengan como ancestros a las morenas, serían presumiblemente no superiores a cero. Es cierto que estoy relacionado a distancia con las morenas, pero el hecho de que encuentre esa estrategia en el medio ambiente del videojuego del laberinto es, casi con toda seguridad, algo que depende de su obvia excelencia, dadas mis necesidades y deseos, y mi propia valoración de mis limitaciones en aquel momento. Este ejemplo ilustra

el obstáculo fundamental —no insuperable pero bastante más importante de lo que se supone— a la inferencia en la sociobiologia humana: demostrar que un cipo concreto de conducta humana es ubicuo o casi ubicuo en culturas humanas separadas ampliamente, no significa probar que exista una predisposición genética para esa conducta concreta. Hasta donde yo sé, en toda cultura conocida por los antropólogos los cazadores arrojan sus lanzas con la punta afilada hacia delante, aunque esto, obviamente, no prueba que exista un gen que diga que la punta afilada debe ir por delante, hasta el extremo de haberse fijado en nuestra especie.

Las especies no humanas pueden exhibir una capacidad similar, aunque reducida, para reinventar la rueda, aun cuando les falte cultura. Los pulpos son notablemente inteligentes y aunque no muestran signos de transmisión cultural, son lo suficientemente inteligentes para que no pueda sorprendernos descubrir en ellos, a título individual, muchas buenas mañas que nunca se les habían planteado como problemas específicos a sus ancestros. Tal uniformidad puede ser interpretada incorrectamente por los biólogos como signos de un «instinto» especial cuando, de hecho, es su inteligencia general la que les conduce, una y otra vez, a encontrar la misma brillante idea. El problema de la interpretación para el *Homo sapiens* se multiplica muchas veces debido a la transmisión cultural. Aunque si algunos individuos cazadores no fueran lo suficientemente brillantes como para darse cuenta por sí mismos de que debían arrojar la lanza con el extremo afilado por delante, les hubiera sido advertido por sus pares, o lo observarían en la práctica, y apreciarían los resultados de inmediato. En otras palabras, si no somos totalmente idiotas, no necesitamos una base genética para aceptar adaptaciones que, de todas formas, acabaríamos aprendiendo de nuestros amigos. Es difícil de creer que los sociobiólogos puedan cometer el error de ignorar esta omnipresente posibilidad, pero la evidencia a favor de que así lo hacen es enorme, una y otra vez (Kitcher 1985). Se podrían comentar muchos ejemplos, aunque deseo concentrarme en un caso particularmente llamativo y bien conocido. Mientras que E. O. Wilson establece claramente que las conductas humanas para ser explicadas por hipótesis genéticas específicas deben ser «lo menos racional del repertorio humano… En otras palabras, deben siempre implicar fenómenos biológicos innatos que son los menos susceptibles de imitación por la cultura» (1978:35), a continuación sigue afirmando, por ejemplo, «que en todas las culturas humanas (a los seres humanos nos gusta reclamar un poco de espacio propio), la evidencia de territorialidad es clara prueba de que, como en muchas otras especies,

tenemos una predisposición genética, con su circuito ya cableado en el momento del nacimiento, para la defensa del territorio» (1978:107ss). Esto puede ser verdad —de hecho, no sería nada sorprendente dado que muchas especies exhiben de modo muy manifiesto una territorialidad innata— y es difícil pensar en una fuerza que pudiera eliminar este dispositivo de nuestra composición genética. Pero la ubicuidad de la territorialidad en la sociedades humanas no es, *por si misma*, evidencia en absoluto de esta base genética, dado que la territorialidad tiene bastante sentido en muchas situaciones humanas. Si la territorialidad no es un movimiento forzado, es algo muy cercano.

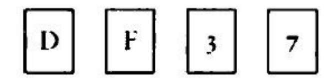
Las mismas consideraciones que en otras partes de la biosfera se tienen en cuenta *para* una explicación, en términos de selección natural, de una adaptación —utilidad manifiesta, valor obvio, razonabilidad innegable del diseño— se tienen en cuenta también *contra la necesidad* de explicación de este tipo en el caso de la conducta humana. Si una maña o habilidad es buena, en ese caso será descubierta de forma rutinaria por toda cultura, sin necesidad de descenso genético o de transmisión cultural de las particularidades**[169]**. En el capítulo 12 vimos que es la perspectiva de una evolución cultural convergente —reinventar la rueda— la que causa estragos a nuestros intentos de convertir la memética en una ciencia. La misma dificultad bloquea, y por idéntica razón, todos los intentos de inferir factores genéticos de hechos culturales comúnmente compartidos. Aunque Wilson se ha dado cuenta algunas veces del problema, en otros momentos lo olvida:

> Las similitudes entre las primeras civilizaciones de Egipto, Mesopotamia, India, China, México y América Central y del Sur en estos grandes temas son muy estrechas. Estas similitudes no pueden explicarse como productos del azar o de la fertilización cultural cruzada (Wilson 1978:89).

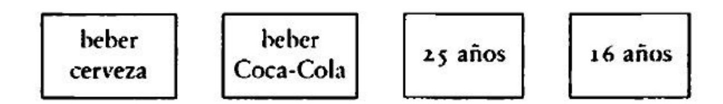
Necesitamos observar sucesivamente cada similitud relevante para ver si alguna de ellas *necesita* de una explicación genética, ya que, además de la fertilización cultural cruzada (descendencia cultural) y el azar, existe la posibilidad de la reinvención. Puede haber factores genéticos específicos operando en muchas de estas similitudes, aunque, como Darwin subrayó, las mejores evidencias serán siempre idiosincrasias —homologías peculiares— y supervivencias que ya no son racionales. Los casos más llamativos de este tipo están siendo descubiertos actualmente en el matrimonio de la sociobiología y de la psicología cognitiva, conocido más recientemente con el nombre de psicología evolutiva (Barkow, Cosmides y Tooby 1992). Resaltar un caso singular aporta un útil contraste entre los buenos y los malos usos del

pensamiento darwinista en el estudio de la naturaleza humana, y clarifica la posición acerca de la racionalidad (o simplemente, de la *no estupidez*) que acabamos de presentar.

¿Hasta qué punto somos lógicos los seres humanos? En algunos aspectos muy lógicos, y en otros vergonzosamente inconsistentes. En 1969, el psicólogo Peter Wason diseñó una simple prueba que gente brillante —alumnos de universidades, por ejemplo— hacen bastante mal. Puede intentarlo. Aquí hay cuatro cartas boca arriba, unas con números y otras con letras. Cada carta tiene un número en una cara y una letra en la otra:



Su tarea consiste en comprobar si, en este caso, la siguiente regla tiene algunas excepciones: *Sí una carta tiene una «D» en una cara, debe de tener un «3» en la otra*. En ese caso, ¿a qué cartas necesita darle la vuelta para descubrir si esta regla es verdad? Es triste admitirlo pero, en la mayoría de los experimentos, poco más de la mitad de los estudiantes dieron la respuesta correcta. ¿Cuál es la suya? La respuesta correcta es bastante más obvia si desplazamos muy levemente el contenido (pero no la estructura) del problema. Si usted es el portero en un bar, su trabajo consiste en no permitir que ningún cliente menor de edad (por debajo de veintiún años) beba cerveza. Las cartas contienen información acerca de la edad en una de sus caras, y de lo que el cliente está bebiendo en la otra cara. ¿A qué cartas necesita darles la vuelta?



A la primera y a la última, obviamente, las mismas que en el primer problema. ¿Por qué es este segundo planteamiento mucho más fácil que el anterior? Podría pensarse en el carácter abstracto del primero frente a la concreción o familiaridad del segundo, o en el hecho de que el segundo implica una regla convencional, no una regularidad de la naturaleza. Se han propuesto, literalmente, cientos de pruebas de Wason a individuos, en cientos de variaciones, probando éstas y otras hipótesis. La realización de las pruebas

por parte de los individuos varía ampliamente, dependiendo de los detalles concretos de la prueba y de sus circunstancias, pero una revisión de los resultados no deja lugar a dudas de que hay propuestas que resultan difíciles para casi todos los grupos de individuos, y otras que son fáciles *para los mismos sujetos*. Pero persiste un enigma, reminiscencia del enigma de las dos cajas negras: ¿qué era exactamente lo que había en los casos difíciles que los hacía difíciles?, o —una pregunta mejor— ¿qué es lo que había en los casos fáciles que los hacía fáciles? Cosmides y Tooby (por ejemplo, Barkow, Cosmides y Tooby 1991:cap. 2) presentaron una hipótesis evolucionista, y es difícil imaginar que esta idea se le ocurriera a cualquiera que no fuera precisamente consciente de las posibilidades del pensamiento darwiniano: los casos fáciles son todos aquellos casos que son interpretados sin dificultad como tareas de vigilancia de un contrato social o, en otras palabras, el descubrimiento de los tramposos.

¡Cosmides y Tooby parecen haber descubierto un fósil de nuestro pasado nietzscheano! Diseñar las hipótesis no significa, naturalmente, probarlas, pero una de las importantes virtudes de la hipótesis de Cosmides y Tooby es que puede ser sometida a unas pruebas muy claras y, hasta ahora, ha resistido muy bien frente a una amplia variedad de intentos de refutación. Supongamos que la hipótesis es verdad; ¿demostraría en tal caso que podemos razonar tan sólo sobre aquellas cosas para las cuales la madre naturaleza nos ha cableado para que podamos razonar? Obviamente no; exactamente muestra por qué es más fácil para nosotros (más «natural») razonar sobre unos temas que sobre otros. Hemos diseñado artefactos culturales (sistema de lógica formal, estadísticas, teoría de la decisión y otros, enseñados en los cursos universitarios) que expanden nuestros poderes de razonamiento al multiplicarlos muchas veces. Incluso los expertos olvidan a veces estas técnicas especializadas y recurren a viejos métodos de razonamiento, técnicamente poco precisos, a menudo con resultados desconcertantes, como muestra la prueba de Wason. Independientemente de cualquier hipótesis darwiniana, sabemos que, excepto cuando la gente es particularmente consciente del uso de estas técnicas de razonamiento poco precisas, tiende a caer en ilusiones cognitivas. ¿Por qué somos tan susceptibles a *estas* ilusiones? El psicólogo evolucionista responde: por la misma razón que somos susceptibles a las ilusiones ópticas y a otras ilusiones sensoriales; estamos construidos así. La madre naturaleza nos diseñó para resolver un cierto grupo de problemas planteados por el medio ambiente en el cual hemos evolucionado, y cuando surge una solución de bajo coste —

una ganga que resolvería muy bien los problemas más apremiantes, aunque no se pueda generalizar— la tendencia es instalarla.

Cosmides y Tooby denominan a estos módulos «algoritmos darwinianos»; se trata de mecanismos como el de la máquina *two-bitser*y sólo que más imaginativos. Nosotros, obviamente, no funcionamos con uno solo de tales mecanismos de razonamiento. Cosmides y Tooby han ido reuniendo evidencias acerca de otros algoritmos para propósitos especiales, útiles en pensamientos acerca de amenazas y otros intercambios sociales, así como otros problemas ubicuos: riesgos, objetos rígidos y contagio. En lugar de disponer de una sola máquina central de razonamiento de uso general, tenemos una colección de artilugios, todos muy buenas —o, al menos, muy buenos en los medios ambientes en los que han evolucionado— y fácilmente convertibles para nuevos propósitos. Nuestras mentes son como las navajas del ejército suizo, dice Cosmides. Una y otra vez descubrimos curiosos vacíos en nuestra competencia, extraños lapsus que nos aportan claves acerca de la particular historia de I+D que explica la maquinaria que subyace bajo la deslumbrante fachada de la cultura. Esta es, seguramente, la vía correcta de los psicólogos, llevar a cabo ingeniería revertida de la mente humana, aunque siempre atentos al fenómeno QWERTY.

En mi opinión, Cosmides y Tooby han hecho hoy uno de los mejores trabajos de psicología darwiniana, y esta es la razón por la que he escogido su ejemplo, aunque debo atemperar mi recomendación con alguna crítica constructiva. La ferocidad de los ataques de los partidarios de Chomsky y Gould es asombrosa y, en plena lucha como se encuentran, Cosmides y Tooby también tienden a la caricatura de la oposición y, a veces, se muestran demasiado impacientes al descartar con escepticismo sus argumentos presentándolos como provenientes de la simple territorialidad defensiva de los obsoletos científicos sociales que aún no saben una palabra sobre la evolución. Esto es a menudo, aunque no siempre, lo que sucede. Incluso si estuvieran en lo cierto al afirmar que esta racionalidad que poseemos como seres humanos es el producto de las actividades de un conjunto de artilugios para propósitos especiales, diseñados por la selección natural, no se puede deducir de esta premisa que nuestra «navaja del ejército suizo» no pueda haber sido usada, una y otra vez, para reinventar la rueda. En otras palabras, todavía ha de demostrarse que cualquier adaptación particular *no* es un producto cultural que responde muy directa y racionalmente a condiciones totalmente recientes. Como Cosmides y Tooby conocen estos aspectos de la

cuestión, evitan cuidadosamente la trampa en la que hemos visto caer a E. O. Wilson, aunque en el calor de la batalla, a veces los olvidan.

Como le sucedió a Darwin, que pasó por alto la inocua posibilidad de la extinción súbita debido a su empeño por esquivar el catastrofismo, Tooby y Cosmides, y los otros psicólogos evolucionistas, tienden a pasar por alto la insulsa posibilidad del descubrimiento independiente de movimientos forzados debido a su obcecación por sustituir el «modelo estándar de la ciencia social» por un modelo apropiadamente darwiniano de la mente. El modelo estándar de la ciencia social contiene entre sus preceptos:

Mientras los animales están rígidamente controlados por su biología, los seres humanos están determinados por la cultura, un sistema autónomo de símbolos y valores. Libres de limitaciones biológicas, las culturas pueden variar arbitrariamente de una a otra y sin límite… El aprendizaje es un proceso con un propósito general, utilizado en todos los dominios del conocimiento (Pinker, 1994:406; véase también Tooby y Cosmides 1992:24-48).

Esto es, sin duda, erróneo por los cuatro costados. Pero, compárese la cita anterior con mi modelo de la ciencia social sólo ligeramente *no* estándar:

Mientras los animales están rígidamente controlados por su biología, la conducta humana está *ampliamente* determinada por la cultura, un *amplio* sistema autómomo de símbolos y valores que, aunque crece desde una base biológica, lo hace aportándose indefinidamente de ella. *Capaces de sobrepasar o escapar* de las limitaciones biológicas en muchos aspectos, las culturas pueden variar de una a otra hasta el punto de que se puedan explicar así muchas de tales variaciones… Aprender *no* es un proceso que valga para todo, pero los seres humanos tienen tantos artilugios que valen para todo, y aprenden a sacar provecho de ellos con tal versatilidad, que la capacidad de aprender, *a menudo*, puede ser tratada como si fuera un regalo de no estupidez enteramente neutral respecto al medio y neutral también respecto al contenido.

Este es el modelo a favor del cual yo he argumentado en este libro; no es una defensa de los ganchos celestes; simplemente reconoce que ahora tenemos grúas de mayor potencia general que las grúas de cualquier otra especie[170]. Como en cualquier campo de la ciencia, en la sociobiología y en la psicología evolucionista, lo mismo se ha hecho mucho trabajo bueno como mucho trabajo malo. ¿Alguna parte de este trabajo representa el mal? Parte de este trabajo es, cuando menos, desalentadoramente descuidado por el mal uso que de él hicieron ideólogos de una convicción o de otra. Pero, en este punto, de nuevo, la escalada de acusaciones produce, como suele ser típico, más acaloramiento que luz. Un ejemplo puede ser representativo de todo este lamentable campo de batalla. ¿Hay patos violadores? Los sociobiólogos han descubierto un patrón de comportamiento común en el cual los machos de algunas especies —como los patos— se aparean violentamente con hembras que obviamente los rechazan. Han denominado a este comportamiento

violación, y este término ha sido censurado por los críticos, más vigorosamente por la bióloga feminista Anne Fausto-Sterling (1985).

Ella tiene en parte razón. Ya dije que no podíamos considerar el hecho de matar a una de las crías gemelas como un «asesinato», dado que no saben lo que están haciendo. Matan pero no cometen un «asesinato». Es imposible para un pájaro *asesinar* a otro pájaro; la palabra «asesinato» se reserva para la muerte intencionada, deliberada, inicua, de un ser humano por otro. (Podemos matar a un oso, pero no asesinarlo y si el oso nos mata no es un asesinato). Entonces, ¿puede un pato *violar* a otro pato? Anne Fausto-Sterling y otras feministas dicen que no, que ésta es una incorrecta aplicación de un término que sólo se aplica apropiadamente a las fechorías humanas. Si hubiera un término común para «violar» —como «matar» es válido tanto para «homicidio» o muerte no premeditada como para «asesinato» o muerte premeditada— en ese caso la utilización por los sociobiólogos del término «violación» para la copulación forzada entre no humanos, en lugar de usar un término con menos carga emocional, sería verdaderamente excesiva. Pero no existe tal término.

Así que, ¿es el uso del breve y descriptivo término «violación», en lugar de la expresión «copulación forzada» (o cualquier otra) un grave pecado? No es, desde luego, lo más apropiado. Pero ¿se quejan los críticos de otros términos extraídos de la vida humana, de uso común entre los sociobiólogos? Existe «canibalismo» sexual entre las arañas (las hembras esperan a que el macho haya terminado de fecundarlas y entonces lo matan y se lo comen), hay gaviotas «lesbianas» (parejas de hembras que permanecen apareadas durante varias estaciones, defendiendo su territorio, construyendo un nido y compartiendo la tarea de incubar los huevos). Hay gusanos «homosexuales» y pájaros «cornudos». Al menos un crítico. Jane Lancaster [1975], ha objetado de hecho la palabra «harén» utilizada para referirse a grupos de hembras guardadas y fecundadas por un solo macho, como sucede con el elefante marino; Jane Lancaster recomienda utilizar la expresión «grupo de un solo macho» dado que las hembras son virtualmente autosuficientes, excepto para la fertilización (Fausto-Sterling 1985:181n). Me parece que el deliberado canibalismo humano es muchísimo más terrible que lo que una araña pueda hacer a otra, pero no tengo objeción si los especialistas en arañas desean utilizar este término. En este sentido, ¿qué decir de los términos benignos (G. Williams 1988)? ¿Pondrán también los críticos objeciones a términos como «ritual del cortejo», la «llamada de alarma», o al uso del vocablo «madre» para referirse a una hembra no humana que tiene una cría o varias crías?

Anne Fausto-Sterling señala que los sociobiólogos que ella critica por utilizar el término violación en los animales, se cuidaban de puntualizar que la violación humana era diferente de la violación en otras especies. Fausto-Sterling cita el siguiente texto de Shields y Shields [1983]:

> En último término, los hombres pueden violar porque con ello aumentan su capacidad de adaptación al medio, y así la violación puede servir, al menos en parte, a una función reproductora, pero en un sentido inmediato es posible que violen porque están irritados u hostiles, como sugieren las feministas (Fausto-Sterling 1985:193).

Este pasaje no es la resonante denuncia de la violación que Fausto-Sterling exige —algo que se da por sobreentendido en el contexto de un artículo científico— pero en él se disocia claramente la violación humana de la «justificación» biológica. Esto hace que Anne Fausto-Sterling se revuelva afrentada y haga responsable a los sociobiólogos de varias declaraciones hechas por abogados de la defensa en casos de violación que, a favor de sus clientes, han asegurado una «insoportable urgencia física» o que han descrito la agresión sexual como «una violación relativamente suave». ¿Qué tienen que ver estas afirmaciones con la sociobiología en general o con los artículos que Fausto-Sterling discute en particular? Fausto-Sterling no ofrece pruebas para creer que estos abogados citen a los socio-biólogos como autoridades, en apoyo de su defensa, ni que conozcan siquiera su existencia. Con igual justicia Fausto-Sterling podría responsabilizar a la tribu de los especialistas en Shakespeare por tales extravíos de la justicia (en el supuesto de que los hubiera) ya que en sus trabajos no han condenado suficientemente el retrato, a veces tolerante, que hace Shakespeare de la violación en sus obras. Este no es el camino para fomentar los análisis clarificadores de estos temas. La irritación se acrecienta y los temas son tremendamente serios, razón por la cual los científicos y los filósofos deben ser cuidadosos, para no abusar ni de la verdad ni de los otros, en nombre de una causa digna.

¿Cuál sería, entonces, un planteamiento más positivo para una ética «naturalizada»? Tengo unas sugerencias preliminares que ofrecer en el próximo capítulo.

CAPÍTULO 17
Un nuevo diseño de la moral

I. *¿Puede ser naturalizada la ética?*[171]

> De este modo, al final el hombre llega a darse cuenta, a través de hábitos adquiridos y quizás heredados, de que es mejor para él obedecer a sus más persistentes impulsos. La imperiosa palabra debe parecer simplemente implicar la consciencia de la existencia de una regla de conducta, sin que importe dónde ésta haya podido originarse.
>
> Charles Darwin, *El origen del hombre*

La cultura humana, la religión en particular, es un depósito de preceptos éticos, que abarca desde el «Trata a tu prójimo como a ti mismo», los diez mandamientos, y el «Conócete a ti mismo» de los griegos, hasta todo tipo de órdenes específicas, prohibiciones, tabús y rituales. Desde Platón, los filósofos han intentado organizar estos imperativos en un sistema universal de ética, racional y defendible, sin que hasta ahora se haya llegado a algo parecido a un consenso. La matemática y la física son las mismas para todos y en cualquier lugar, pero la ética todavía no ha podido alcanzar un similar y ponderado equilibrio[172]. ¿Por qué no? ¿Es ilusorio este objetivo? ¿Es la moral un asunto de gusto subjetivo (y de poder político)? ¿No existen verdades éticas que puedan ser descubiertas y confirmadas, que no sean movimientos forzados o buenas mañas en el proceso evolutivo? Se han construido grandes edificios de teoría ética que han sido criticados y defendidos, sometidos a revisión y ampliados por los mejores métodos de la investigación racional, y entre estos artefactos del razonamiento humano figuran algunas de las creaciones más extraordinarias de la cultura que, sin embargo, todavía siguen sin suscitar el consenso incondicional de todos aquellos que las han estudiado cuidadosamente.

Quizás podamos descubrir algunas claves acerca del estado y las perspectivas de la teoría ética si reflexionamos sobre las limitaciones del gran proceso de diseño, entre los cuales se encuentran hasta hoy día quienes se ocupan de la ética. ¿Qué se deduce, podemos preguntar, del hecho de que las

tomas de decisiones éticas, como todo proceso real de exploración en el espacio de diseño, sea, hasta cierto punto, un proceso miópico y bajo la presión del tiempo?

Poco después de la publicación del libro de Darwin, *El origen de las especies*, otro eminente Victoriano, John Stuart Mill, publicó su propuesta de una teoría ética universal en un libro titulado *Utilitarism* (1861). Darwin lo leyó con interés y respondió al «celebrado trabajo» en su libro *El origen del hombre* (1871). Darwin quedó perplejo ante la posición de Mili sobre si el sentimiento moral era innato o adquirido, y requirió la ayuda de su hijo William, quien advirtió a su padre que Mili se encontraba «más bien confundido con todo el tema» (R. Richards 1987:209n), pero, aparte de unos puntos de desacuerdo, Darwin y Mili fueron considerados (correctamente) como postulantes del naturalismo, y por ello reprendidos por los defensores de los ganchos celestes, especialmente por St. George Mivart, quien declaró:

Los hombres tienen conciencia de una absoluta e inmutable regla que requiere, legítimamente, obediencia a una autoridad necesariamente suprema y absoluta; en otras palabras, se construyen juicios intelectuales que implican, en la mente que juzga, la existencia de un ideal ético (Mivart 1871:79).

Ante tal ventolera no hay, probablemente, mejor respuesta que la cita de Darwin que encabeza este capítulo. Pero hubo, también, críticas más mesuradas y una de las más frecuentes considera desagradable e inaceptable la posición de Mili: «Los mismos defensores de la utilidad se encuentran, a menudo, obligados a replicar a tales objeciones diciendo que no existe un tiempo, previo a la acción, para calcular y sopesar los efectos de una línea de conducta basada en la felicidad general». La reacción de Mili fue bastante agresiva:

Realmente, los hombres deben dejar de decir desatinos sobre este tema, del cual ni hablarían ni escucharían en otros asuntos de interés práctico. Nadie sostiene que el arte de la navegación no esté fundamentado en la astronomía por el hecho de que los marinos lo calculen en el almanaque náutico. Al ser criaturas racionales, los marinos salen a navegar con los cálculos ya hechos; y todas las criaturas racionales se lanzan al mar de la vida con sus mentes preparadas para las preguntas habituales acerca de lo correcto y lo erróneo, así como para muchas de las aún más difíciles cuestiones sobre lo juicioso y lo insensato. Y esto, mientras la previsión continúe siendo una cualidad humana, seguirá haciéndose (Mili 1861:31).

Esta orgullosa respuesta ha sido bien acogida por muchos teóricos de la ética —quizá demasiados—, pero de hecho lo que hace es encubrir, sin cerrarla, una grieta que ha sido gradualmente ensanchada por el asalto de los críticos. Los objetores partían de una curiosa premisa errónea al pensar que un sistema de pensamiento ético *debía supuestamente funcionar* y observaron que el

sistema de Mili era muy difícil de llevar a la práctica. Mili insistía en que no había objeción: se supone que el utilitarismo es práctico, pero no hasta ese punto de aplicación práctica. Su verdadero papel consiste en ser un trasfondo justificador de los hábitos de pensamiento de los verdaderos razonadores morales, que se encuentran en el primer plano. Este papel de trasfondo adjudicado a la teoría ética (al que no solo ha aspirado el utilitarismo) ha probado, no obstante, estar mal definido y ser inestable. ¿Cómo se supone que pueda ser práctico un sistema de pensamiento ético? ¿Para qué sirve una teoría ética? Diferencias tácitas de opinión acerca de este tema, e incluso un cierto grado de falsa conciencia entre los protagonistas, se han añadido a la falta de conclusiones del subsiguiente debate.

En su mayoría, los filósofos se han complacido en ignorar los problemas prácticos de la toma de decisiones en tiempo real, haciendo referencia al simple hecho de que todos somos finitos y olvidadizos, y que tenemos que apresurarnos a emitir juicios, como un elemento de fricción, real pero irrelevante, en la maquinaria que ellos tratan de describir. Es como si se tratara de dos disciplinas: la ética propiamente dicha, que tiene a su cargo la tarea de calcular los principios que determinan lo que el agente ideal debe hacer bajo cualquier circunstancia y, al mismo tiempo, la disciplina menos interesante y «simplemente práctica» de la «moral entendida como primeros auxilios» o, dicho de otro modo, «¿Qué hacer hasta que el doctor de filosofía llegue?»; esta segunda disciplina tratará, en términos simples y directos, de cómo tomar decisiones sobre la marcha bajo la presión del tiempo.

En la práctica, los filósofos lo saben, pasamos por alto importantes consideraciones —consideraciones que realmente no deberían pasar inadvertidas— y nuestro pensamiento se ve influido por cientos de modos peculiares, y moralmente indefendibles; pero en principio, lo que debemos hacer es lo que la teoría ideal (una teoría ideal u otra) dice que debemos hacer. Los filósofos han concentrado sus esfuerzos, de manera prudente, en exponer lo que es la teoría ideal. No pueden negarse los frutos teóricos de una deliberada hipersimplificación a través de la idealización, en filosofía o en cualquiera disciplina científica. La realidad, en toda su desordenada particularidad, es demasiado complicada para teorizar sobre ella de una manera directa. El asunto consiste más bien en saber (dado que toda idealización es una elección estratégica) qué idealización puede realmente arrojar alguna luz sobre la naturaleza de la moral, y cuál nos obliga a aceptar entretenidos cuentos de hadas.

Es fácil olvidar lo poco prácticas que realmente son las teorías éticas, pero podemos hacer que la verdad resplandezca reflexionando sobre lo que está implícito en el uso por parte de Mili de una metáfora extraída de la tecnología de su tiempo. El *Almanaque Náutico* es una especie de efemérides astronómica, un libro de tablas calculadas y publicadas anualmente, de las cuales uno puede, fácil y rápidamente, derivar la exacta posición en los cielos del sol, la luna, los planetas y las estrellas mayores, para cada segundo del próximo año. La precisión y certeza de este generador anual de expectativas era, y aún es, un ejemplo inspirador del poder de la previsión humana, disciplinada adecuadamente por un sistema científico y *aplicada a un tema suficientemente ordenado*. Armado con los frutos de este sistema de pensamiento, el marino racional puede naturalmente aventurarse, confiado en su habilidad para tomar decisiones sobre la navegación, en tiempo real y suficientemente informado. Los métodos prácticos diseñados por los astrónomos funcionan realmente.

¿Disponen los utilitaristas de un producto similar que puedan ofrecer al público en general? J. Stuart Mill parece, al principio, decir que sí. Estamos hoy habituados a los ampulosos discursos pronunciados en nombre de docenas de sistemas de alta tecnología —análisis coste-beneficio, sistemas expertos basados en computación, etc.— y desde la perspectiva actual podemos suponer a Mill implicado en un inspirado papel publicitario cuyo objetivo sería sugerir que el utilitarismo puede aportar al agente moral una infalible «ayuda para la toma de decisiones». («¡Para su uso hemos hecho estos difíciles cálculos! Todo lo que necesita es sencillamente rellenar las cuadrículas del simple formulario que se le ha facilitado»).

Jeremy Bentham, el fundador del utilitarismo, aspiraba ciertamente a este «felicific calculus» que incluirá estribillos mnemotécnicos, al modo de los sistemas de navegación práctica por las estrellas que memoriza todo capitán de un barco.

> Intenso, largo, cierto, veloz, fructífero, puro;
> tales señales en placeres y en dolores perduran,
> tales placeres busca si privado es tu fin:
> si es público, déjalos que se extiendan ampliamente.
>
> Benrham,
> *Introduction to the Principles of Morals and Legislation*

Bentham fue, de buen grado, un reduccionista insaciable —se podría decir que el B. F. Skinner de su tiempo— y este mito de la búsqueda de lo práctico ha sido, desde su inicio, parre de la retórica del utilitarismo. Pero en Stuart

Mill ya podemos ver el inicio de la retirada a la torre de marfil del idealismo, a lo que es calculable «en principio», pero no en la práctica.

La idea de Mili era, por ejemplo, que la mejor de las homilías y reglas prácticas de la moralidad cotidiana —las fórmulas que la gente *considera realmente* en el turbulento curso de sus deliberaciones— había recibido (o recibiría, en principio) apoyo oficial desde el complejo, laborioso y sistemático método del utilitarismo. La fe asentada en estas fórmulas por el agente racional medio, basada como estaba en muchas experiencias vitales acumuladas en la memoria cultural, podría justificarse («en principio») deduciéndola formalmente de la teoría. Pero tal deducción jamás se ha conseguido**[173]**.

La razón no es difícil de entender: es extremadamente improbable que pudiera haber un *algoritmo factible* para el tipo de análisis global del coste-beneficio que requiere el utilitarismo (o cualquiera otra teoría «consecuencialista»). ¿Por qué? Debido a lo que puede denominarse el efecto «Three Mile Island». ¿Fue la fundición de las varillas que contienen el combustible nuclear, tal como ocurrió en la planta de Three Mile Island, una cosa buena o mala? Si, al planificar el curso de algunos acontecimientos, encontráramos esta fundición como una secuela de probabilidad *p*, ¿qué «peso» le asignaríamos? ¿Es un resultado negativo que debemos evitar o un resultado positivo que debe ser cuidadosamente fomentado**[174]**? No podemos decirlo todavía y no está claro que en *ningún* plazo concreto podamos dar la respuesta. (Nótese que éste no es un problema de medidas insuficientemente precisas; lo que sucede es que no podemos determinar *el signo*, positivo o negativo, del valor que se le asigne al resultado).

Comparemos el problema que aquí consideramos con los problemas que afrontan los diseñadores de programas de ajedrez. Cabe suponer que el método para responder al problema de la presión del tiempo real en las técnicas de tomas de decisión éticas es similar al modo en que uno responde en el ajedrez a la presión del tiempo: las técnicas heurísticas en las que se van buscando —en el árbol de la toma de decisiones— aquellas opciones que se decide descartar. Pero en la vida no existe el jaque mate, no existe un punto en el cual consigamos un resultado definitivo, sea positivo o negativo, desde el cual podamos calcular, mediante análisis retrospectivo, los valores reales de las alternativas que se encuentran a lo largo de los caminos que hemos elegido. ¿Hasta qué punto debe ser profundo el análisis antes de decidir en el ajedrez acerca del «valoración» de una posición? En el ajedrez, un análisis que parece positivo para un jugador que juega contra un programa de

ordenador con un *ply* (número de movimientos que el ordenador controla por adelantado) de 5, puede parecer desastroso con un *ply* de 7. Hay métodos para afinar los procedimientos heurísticos de búsqueda de modo que se minimice (aunque no definitivamente) el problema provocado por la evaluación errónea de los movimientos anticipados. ¿Es la captura anticipada un futuro suficientemente positivo para que sea nuestro objetivo o bien puede ser el comienzo de un brillante sacrificio para nuestro contrincante? Un *principio de reposo* ayudará a resolver esta cuestión: examinemos siempre unos cuantos movimientos previamente a cualquier serie rápida de intercambios de piezas, con el fin de observar el aspecto del tablero cuando está en reposo. Tras el accidente de la Three Mile Island, ha seguido más de una década de consolidación y reposo (sucedió en 1980), pero, después de sopesar todos los aspectos de la cuestión, aún no tenemos idea de si este accidente debe contarse entre las cosas buenas que han sucedido o entre las malas.

La sospecha de que no existe una solución estable y persuasiva para tales callejones sin salida subyace bajo la superficie de críticas al «consecuencialismo», el cual aparece ante muchos escépticos como una versión levemente velada del vacío consejo frecuentemente oído en el mercado de valores: «Compra barato y vende caro», una gran idea, en principio, pero sistemáticamente inútil como consejo a seguir**[175]**.

De modo que no sólo los utilitaristas no han llevado nunca a la práctica real la determinación de sus elecciones morales específicas mediante el cálculo de las utilidades esperadas de (todas) las alternativas (no ha habido tiempo, como hacía notar nuestro original objetor), sino que nunca han conseguido derivaciones estables, fuera de tiempo real, de resultados parciales —«hitos o señales de dirección», como señaló Mili— para ser aprovechados en el momento por aquellos que deben enfrentarse con «asuntos de consecuencias prácticas».

Entonces, ¿qué sucede con los principales rivales de los utilitaristas, como las distintas variedades de kantianos? Del mismo modo, la retórica de los kantianos ha pagado tributo a la «practicabilidad», en gran parte a través de sus acusaciones sobre la «*im*-practicabilidad» de los utilitaristas**[176]**. ¿Qué proponen los kantianos para sustituir los impracticables cálculos «consecuencialistas»? El seguimiento de máximas, actitud derivada, a menudo, de una veneración por las reglas, de un tipo o de otro, tal como la invocada en una de las formulaciones de Kant [1785] del *imperativo categórico*: Actúa sólo movido por aquella máxima que al mismo tiempo querrías que llegara a ser una ley universal. La toma de decisiones kantiana

revela, de modo característico, diferentes idealizaciones —huidas de la realidad en otras direcciones— que hacen todo el trabajo. Por ejemplo, a menos que dispongamos de un *deus ex machina*, un accesible maestro de ceremonias que nos susurre sugerencias al oído, es difícil imaginar cómo podríamos hallar el modo de limitar el propósito de las «máximas» referidas a nuestras acciones potenciales antes de pasarlas por la «prueba del tornasol» del imperativo categórico. Parece existir una inacabable provisión de máximas potenciales.

La peregrina esperanza de Jeremy Bentham de hallar un procedimiento de decisión ante los problemas éticos, consistente en rellenar una hoja con una serie de casillas en blanco, es ciertamente tan extraña al espíritu del moderno Kant como lo es a los sofisticados utilitaristas. Todos los filósofos pueden estar de acuerdo, según parece, en que el pensamiento moral real requiere visión e imaginación, y no puede alcanzarse mediante la aplicación de fórmulas. Como señaló el propio Mili (1871:31) incluso con gran enojo: No es difícil probar que cualquier criterio ético es insatisfactorio si lo suponemos imbuido en la idiotez universal. Sin embargo, este brote de retórica está en conflicto, de alguna manera, con una analogía previa del propio Mili, ya que una de las alegaciones legítimas de los sistemas de navegación práctica era, precisamente, que cualquier idiota podía manejarlos.

No pretendo, en modo alguno, que esto se convierta en una desagradable denuncia, sino en el recordatorio de algo completamente obvio: ningún sistema de ética remotamente preciso ha podido ser transformado en *tratable computacionalmente*, ni siquiera indirectamente, para problemas morales del mundo real. De este modo, aunque no puede decirse que haya habido carestía de *argumentos* utilitaristas (y kantianos, etc.) a favor de políticas particulares, instituciones, prácticas y actos, éstos han sido todos protegidos fuertemente con condiciones *ceteris paribus* y alegatos de verosimilitud acerca de sus idealizantes hipótesis. Estas protecciones se diseñan para superar la amenaza de explosión combinatoria de cálculos si realmente intentamos *considerar todos los factores*, tal como nos dice la teoría. Y como argumentos —no como deducciones— todos han sido controvertidos (lo que no significa que ninguno de ellos posea fundamento en el análisis final).

Para obtener una mejor visión de las dificultades que contribuyen al razonamiento moral *real*, nos plantearemos un pequeño problema moral y veremos qué podemos hacer. Aunque alguno de sus detalles son exóticos, el problema que planteo ejemplifica una estructura que resulta familiar.

2. *Evaluación de candidatos*

Nuestro departamento de filosofía ha sido escogido para administrar un magnífico legado: una beca de doce años de duración en abierta competición con los más prometedores graduados en Filosofía del país. El departamento anuncia debidamente la beca y sus condiciones en el *Journal of Philosophy* y al finalizar el plazo de presentación de solicitudes quedamos decepcionados al recibir 250.000 inscripciones legales, complementadas con gruesos volúmenes de documentación de los candidatos, ejemplares de las publicaciones y escritos de presentación de los aspirantes. Un rápido cálculo nos convence de que, si se asume la obligación de evaluar todo el material de los candidatos inscritos antes del cierre de la convocatoria, no sólo se impediría que el departamento cumpla su misión principal de enseñar, sino que —dados los costes de la administración y contratación de evaluadores cualificados adicionales— se provocaría la bancarrota del propio legado, de modo que todo se gastaría en la evaluación: nadie saldría ganando.

¿Qué hacer? Si hubiéramos previsto este enorme número de candidatos, podríamos haber impuesto requisitos para la beca mucho más restringidos, pero ahora ya es tarde: cada uno de los 250.000 candidatos tiene, cabe suponer, igual derecho a ser considerado y, al llegar al acuerdo de administrar la competición, hemos de aceptar la obligación de seleccionar el mejor candidato. (No trato de dar por sentada cualquier cuestión con esta formulación en términos de derechos y obligaciones. Si esto le importa a alguien, replanteamos la situación del problema en términos de la inconveniencia global de violar las condiciones dadas a conocer cuando se anunció la competición. Mi opinión es que nos pondríamos en un aprieto, cualquiera que fuese nuestra convicción ética). Antes de seguir leyendo, por favor, tómese el tiempo que considere conveniente para planificar su propia solución al problema (¡por favor, nada de fantasías sobre soluciones tecnológicas!).

Cuando he planteado este problema a colegas y estudiantes, he encontrado que, tras un breve período de análisis, tienden a asentarse en una u otra versión de una estrategia mixta, tales como:

1) Escoger un pequeño número de criterios de excelencia fácilmente comprobables y no enteramente asintomáticos, tales como puntuación media en la graduación, número de cursos de filosofía completados, peso del *dossier* (eliminando el demasiado liviano y el demasiado pesado); utilizamos estos criterios para un primer corte.

2) Realizar un sorteo de lotería con los restantes candidatos, disminuyendo el número de éstos de manera aleatoria hasta un número final de finalistas manejable, entre cincuenta o cien.

3) Estos *dossiers* serán cuidadosamente revisados por un comité, el cual votará el ganador.

No hay duda de que es muy improbable que este procedimiento encuentre al mejor candidato. De hecho, las probabilidades están a favor de que algo más de unos pocos de los perdedores, si van a un juicio, podrían convencer a un jurado de que ellos eran obviamente superiores al elegido como ganador. Pero alguien podría replicar diciendo que esto es difícil; lo hicimos lo mejor que pudimos. Es muy posible, naturalmente, que perdiéramos el juicio, aunque seguiríamos pensando, correctamente, que en aquel momento habíamos llegado a la mejor solución.

Mi ejemplo trata de ilustrar, de forma amplificada y a cámara lenta, las características ubicuas de la toma de decisión en tiempo real. En primer lugar, existe la simple imposibilidad física de «tener en cuenta todos los factores» en el tiempo asignado. Nótese que «todos los factores» no tienen que significar *todas y cada una de las cosas*, ni siquiera todas *las personas de este mundo*, sino únicamente *todo lo que hay en 250.000 dossiers fácilmente disponibles*. Tenemos a la mano toda la información que necesitamos; en este ejemplo no hay necesidad de plantear la realización de investigaciones adicionales. En segundo lugar, está el uso implacable y perentorio de algunas reglas claras de corte secundarias. Nadie piensa que la nota media de la graduación sea un indicador remoto e infalible de una promesa, aunque probablemente sea algo superior al peso del *dossier*, y evidentemente superior al número de letras en el nombre del candidato. Hay un cierto toma y daca entre la facilidad de aplicación y la fiabilidad, y si uno no puede pensar *rápidamente* sobre algunos criterios fácilmente aplicables, en los cuales se pueda tener *alguna* fe, lo mejor sería eliminar el paso 1) y proceder directamente a la lotería para todos los candidatos. En tercer lugar, la lotería ilustra una abdicación parcial del control, dejando de lado una parte de la tarea y permitiendo a otra instancia —naturaleza o suerte— hacerse cargo durante algún tiempo, aunque sin dejar de asumir la responsabilidad por los resultados. (Esta es la parte alarmante). En cuarto lugar, hay una fase en la que tratamos de salvar algo presentable del resultado de este proceso salvaje; tras haber *hiper*simplificado la tarea, se cuenta con un proceso de automonitorización para corregir o renormalizar o mejorar, de algún modo, el producto final. En quinto lugar está la vulnerabilidad inacabable del proceso ante las conjeturas y las revisiones

retrospectivas de otros acerca de lo que hemos hecho, aunque ¡a lo hecho, pecho! Dejemos el resultado como está y pasemos a otra cosa. La vida es corta.

El proceso de toma de decisión que hemos descrito es un ejemplo del patrón fundamental analizado explícitamente en primer lugar por Herbert A. Simon (1975, 1959) y denominado *satisficing*. (En contraste con la *optimización* —obtener el mejor resultado— se trata de obtener un resultado que sea bastante bueno, es decir, que sea *satisfactorio*,) Nótese cómo el patrón se repite a sí mismo, casi como una curva fractal, cuando seguimos el camino descendente de las subdecisiones, las sub-sub-decisiones, etc., hasta que el proceso se hace invisible. En la reunión del departamento convocada para analizar el tratamiento de este dilema, a) todo el mundo estalló con sugerencias, más de las que podían ser razonablemente discutidas en las dos horas permitidas, así que b) el director del departamento decidió ser terminante y no dar la palabra a varios miembros que bien podían, naturalmente, haber aportado algunas buenas ideas y entonces c) después de una breve «discusión» abierta —para que todos pudieran decir lo que quisieran— en la cual el volumen y timbre de voz pueden tener más importancia que el contenido, d) el director del departamento intentó hacer un resumen eligiendo unos cuantos puntos relevantes que, de algún modo, le llamaron la atención como puntos operativos, y la fuerza y la debilidad de éstos se debatió de una manera más ordenada y al final se votó. Después de la reunión, e) había quienes aún pensaban que se podían haber escogido mejores reglas para el corte de los candidatos, que el departamento podía haber proporcionado el tiempo necesario para evaluar doscientos finalistas (o haber restringido el número a veinte), etc., pero, ¡a lo hecho, pecho! Habían aprendido la importante lección de cómo vivir con la toma de decisiones en condiciones subóptimas de sus colegas, así que después de unos minutos u horas de exuberancia en inteligentes discusiones retrospectivas, abandonaron el asunto.

Pero ¿por qué *debemos* abandonar este asunto?, se preguntó a sí mismo uno de los asistentes a la reunión del departamento, pregunta que volvió a hacerse en medio de la discusión, cuando el director del departamento no le dio la palabra. A este profesor le hervía la cabeza a) con las razones por las que debía insistir en ser escuchado, compitiendo con calma con sus colegas, y con las razones por las cuales deberíamos seguir adelante, y todos competían con sus intentos por seguir el hilo de lo que otros estaban diciendo, y así sucesivamente; más información a la mano de la que se puede manejar, así

que b) rápida, arbitraria e irreflexivamente bloquea parte de esa información; corriendo el riesgo de ignorar las consideraciones más importantes y entonces c) se rinde tratando de *controlar* sus pensamientos; renuncia al metacontrol y permite que sus pensamientos puedan conducirle durante algún tiempo a otro lugar. Después de un breve intervalo, este profesor d) recupera el control, intenta alguna ordenación y mejoramiento de los materiales que han sido vomitados durante la discusión y toma la decisión de abandonar el asunto, sufriendo e) súbitas punzadas de duda y de remordimiento pero, como es inteligente, también se encoge de hombros.

Y ¿cómo llega ese profesor a descartar la evanescente e inarticulada microinterrogación («¿Por qué debemos abandonar este tema?»). Aquí el proceso se hace invisible al ojo desnudo de la introspección, pero si observamos los modelos de la «toma de decisiones» y de la «solución de problemas» en la ciencia cognitiva, *dentro* de estos rápidos e inconscientes procesos, como son la percepción y la comprensión del lenguaje, observamos atractivas analogías de nuestras fases con los variados modelos de búsqueda heurística y de solución de problemas[177].

Como hemos visto en este libro, una y otra vez, la toma de decisiones bajo la presión del tiempo se produce, como en este caso, en *sentido descendente*. La búsqueda de un resultado satisfactorio se extiende incluso retrospectivamente, por detrás del diseño biológico estable del agente de la toma de decisiones, hasta las «decisiones» sobre diseños que la madre naturaleza establece cuando nos diseña a nosotros y a otros organismos. Puede haber líneas divisorias, en cierto modo no arbitrarias, trazadas entre las manifestaciones biológicas, psicológicas y culturales de esta estructura, pero no sólo son las estructuras —y sus poderes y sus vulnerabilidades— básicamente las mismas; los contenidos particulares de la «deliberación» no están probablemente encerrados en un nivel específico de la totalidad del proceso, sino que pueden migrar. Bajo la provocación apropiada, por ejemplo, uno puede recuperar alguna consideración virtualmente subliminal y elevarla para su formulación y apreciación autoconsciente —llegando a ser una «intuición»— y expresarla entonces de modo que otros puedan también considerarla. Moviéndose en la otra dirección, una razón para actuar, continuamente mencionada y debatida en comité, puede eventualmente «pasar como algo sobreentendido» —al menos en voz alta— aunque continúa conformando el pensamiento, tanto del grupo como de los individuos, desde alguna base (o bases) más subliminal de operaciones en el proceso. Como han argumentado Donald Campbell [1975] y Richard Dawkins [1976:cap. 11], las

instituciones culturales pueden, a veces, ser interpretadas como compensaciones o correcciones de las «decisiones» tomadas por selección natural.

Los fundamentos que conducen a optar por una solución «bastante buena» —satisfactoria, hecho que constituye la estructura *básica* de todo proceso de toma de decisiones real, sea moral, prudencial, económico e incluso evolucionista— dan origen a una conocida y preocupante inasibilidad de la afirmación que hace confusa la teoría en varias escuelas. Para comenzar, téngase en cuenta que afirmar simplemente que esta estructura es básica no equivale necesariamente a decir que es la mejor, aunque esta conclusión es ciertamente incitada e incitante. Recuérdese que hemos iniciado esta exploración analizando un *problema* moral y tratando de *resolverlo*: el problema de diseñar un *buen* proceso de evaluación de candidatos (un proceso justificado, defendible, bien fundamentado). Supongamos que decidimos que el sistema que hemos diseñado es el mejor posible, dadas las limitaciones. Un grupo de agentes aproximadamente racionales —nosotros— decide que ésta es la vía correcta para diseñar el proceso, y que disponemos de razones para escoger los criterios de evaluación que hemos escogido.

Dada esta geneología, podemos tener el descaro de declarar que éste es el diseño óptimo; el mejor de todos los diseños posibles. Esta aparente arrogancia se me podría imputar a mí mismo desde el momento en que plantee el problema, ya que mi propuesta no es examinar cómo *nosotros alguien debe* tomar decisiones morales, sino cómo *tomamos de hecho* una particular decisión moral. ¿Quiénes somos nosotros para marcar el paso? Bueno, ¿y en quién si no vamos a confiar? Si no podemos confiar en nuestro propio buen juicio, me parece que no podremos comenzar.

Así, lo que nosotros pensamos y cómo pensamos, es evidencia de los principios de racionalidad, de lo que pensamos y de cómo debemos pensar. Este es, en sí mismo, un principio metodológico de racionalidad: llamémosle el *principio Factunorm* (del hecho a la norma). Aceptamos (implícitamente) el principio Factunorm cada vez que intentamos determinar qué o cómo debemos pensar. Porque debemos, en este concreto intento, pensar. Y a menos que podamos creer que lo que pensamos, y cómo lo pensamos en ese momento es correcto —y, por lo tanto, evidencia de lo que pensamos y cómo debemos pensar—, no podremos determinar el qué y cómo debemos pensar (Wertheimer 1974:110-11; véase también Goodman 1965:63).

Sin embargo, las afirmaciones de «optimalidad» tienen un modo de desvanecerse; no es necesario ningún descaro para admitir modestamente que, dadas nuestras limitaciones, aquella era la mejor solución que podíamos encontrar. A veces se comete el error de suponer que hay, o debe haber, una perspectiva única (mejor o más elevada) desde la cual evaluar la racionalidad

ideal. ¿Sufre el agente racional ideal el problema demasiado humano de no ser capaz de recordar ciertas consideraciones cruciales cuando éstas serían las más reveladoras y efectivas para resolver un dilema? Si estipulamos, como simplificación teórica, que nuestro imaginado agente ideal es inmune a tales desórdenes, entonces no llegaremos a plantear la cuestión de cuál sería el modo ideal para afrontarlos.

Tales ejercicios presuponen que ciertas características —las «limitaciones»— son fijas y otras son maleables; la función de las segundas consiste en ser maleables para ajustarse mejor a las primeras. Pero siempre se puede cambiar la perspectiva y preguntar si una de las características presumiblemente maleables no se encuentra, de hecho, fija en una posición, como una limitación que ha de ser acomodada. Y se puede preguntar acerca de cada una de las características fijas, si sería deseable alterarlas en alguna circunstancia; quizá para mejorar. Plantear esta cuestión requiere considerar aún otras características siguientes como fijas, para juzgar lo acertado de la característica que estamos examinando. No contamos aquí con el apoyo de Arquímedes; si suponemos que los lectores del *Manual de primeros auxilios morales* son unos *completos* idiotas, nuestra tarea es imposible; mientras que si suponemos que son santos, nuestra tarea sería demasiado fácil como para arrojar alguna luz.

Esto se aprecia gráficamente en las resbaladizas hipótesis acerca de la racionalidad en la discusión teórica del Dilema del Prisionero; no hay problema si estamos dispuestos a asumir que los jugadores son santos; después de todo, los santos siempre cooperan. Los pelmazos cortos de vista siempre desertan, así que no hay esperanza para ellos. ¿Qué hace el jugador «idealmente racional»? Quizá, como alguien ha dicho, considera racional adoptar la metaestrategia de convertirse a sí mismo en un jugador menos que idealmente racional, con el fin de enfrentarse con los jugadores que son menos que idealmente racionales, con los que sabe que puede enfrentarse. Pero, entonces, ¿en qué sentido es este nuevo jugador menos que idealmente racional? Es un error suponer que esta inestabilidad puede eliminarse si pensamos bastante cuidadosamente sobre lo que es la racionalidad ideal. Esta es una *verdadera* falacia panglossiana. (Más reflexiones en esta línea de pensamiento se encuentran en Gibbard 1985 y Sturgeon 1985).

3. *El manual de primeros auxilios morales*

Entonces, ¿cómo podemos esperar la regulación, o al menos el mejoramiento de nuestro proceso de toma de decisiones éticas, si es irremediablemente heurístico, presionado por el tiempo y miópico? Considerando paralelamente lo que sucede en la reunión del departamento y lo que sucede en nosotros mismos, podemos ver lo que son los metaproblemas, y cómo pueden tratarse. Necesitamos hábitos de pensamiento «alertas» e «inteligentes» o, en otras palabras, colegas que regularmente, si no infaliblemente, guíen nuestra atención hacia direcciones en las que no nos arrepentiremos de penetrar intelectualmente. No tiene sentido tener más de un colega si se trata de colegas clónicos que desean todos plantear la misma objeción, especialistas cortos de miras preocupados únicamente en proteger un cierto grupo de intereses (Minsky 1985).

Ahora bien, ¿cómo evitaremos la existencia de una cacofonía de colegas? Necesitamos algunos colegas que moderen la conversación. Además de los oportunos y apropiados generadores de consideraciones, son necesarios otros que frenen a los que generan la conversación. Necesitamos algunas tácticas que arbitrariamente den por cerradas las reflexiones y disquisiciones de nuestros colegas y corten el debate, independientemente del contenido específico que se esté discutiendo. ¿Por qué no una *palabra mágica*? Las palabras mágicas funcionan muy bien como dispositivos de control en los programas de inteligencia artificial, pero estamos hablando de controlar a colegas inteligentes y éstos no son susceptibles a las palabras mágicas, como si se encontraran bajo una sugestión posthipnótica. Es decir, los buenos colegas serán reflexivos y racionales y de mentalidad abierta, dentro de los límites impuestos por la estrechez de miras de su especialización. Si los más simples mecanismos de los que estamos compuestos son sistemas balísticos intencionales, como he sostenido en el capítulo anterior, nuestros subsistemas, más sofisticados, como nuestros colegas en la realidad, son sistemas intencionales *guiables indefinidamente*. Necesitan ser impactados con algo que sea atractivo a su racionalidad y que al mismo tiempo les haga desistir de seguir reflexionando.

Esto no significa que toda esta gente tenga que estar filosofando *indefinidamente*, retrotrayéndonos indefinidamente a los primeros principios y exigiendo una justificación para estos principios que aparentemente (y realmente) son completamente arbitrarios. ¿Cómo se podría proteger de tales implacables escrutinios a un moderador arbitrario y, en cierto modo, de segunda categoría? ¿Una metapolítica que prohíba la discusión y reconsideración de los moderadores? Pero, podrían preguntar nuestros

colegas, ¿es *ésta* una política adecuada? ¿Estará justificada? Eso no siempre generaría seguramente los mejores resultados, y… así sucesivamente.

Este es un asunto de delicado equilibrio, en el que se cometen errores por ambas partes. Por un lado, debemos evitar el error de pensar que la solución consiste en *más racionalidad*, más reglas, más justificaciones, ya que no hay fin para este tipo de demandas. Cualquier política *puede* ser cuestionada, así que, si no optamos por alguna ruda y arracional conclusión del tema discutido, terminaremos diseñando un proceso de decisión que se moverá en espiral hasta el infinito. Por otra parte, cualquier simple dato en bruto acerca de cómo estamos construidos es —o debería ser— susceptible de ser descartado en futuras reflexiones**[178]**.

No podemos esperar que en este caso haya una única solución estable para tal problema de diseño sino, más bien, una variedad de equilibrios inciertos y temporales, en la cual los moderadores tenderán a ir superponiendo finas capas de pruebas de apoyo que no podrán ser fácilmente cuestionadas y que en un momento dado servirán, afortunadamente, para bloquear y terminar con las consideraciones.

He aquí algunos prometedores ejemplos:

Pero esto haría más mal que bien.
Pero esto sería un crimen.
Pero esto sería romper una promesa.
Pero esto sería utilizar a alguien simplemente como un medio.
Pero esto violaría el derecho de una persona.

En cierta ocasión Jeremy Bentham rechazó bruscamente la doctrina de los «derechos naturales e imprescriptibles» como «una necedad ampulosa», y podemos replicar ahora que probablemente estaba en lo cierto. Quizás hablar de derechos es «una necedad ampulosa», pero una buena *necedad*; y buena solamente porque es ampulosa debido a que tiene el poder «político» para mantenerse elevada por encima de las metareflexiones —no indefinidamente, pero habitualmente «a bastante altura»—, y así reafirmarse como un obligatorio —esto es, moderador— «principio básico».

Puede parecer que un cierto tipo de «adoración de la regla» es una buena cosa, al menos para agentes diseñados como nosotros. Es una buena cosa, no debido a que haya una cierta regla, o un conjunto de reglas (lo cual es probablemente lo mejor, o lo que siempre genera la respuesta correcta) sino porque tener reglas funciona —de algún modo—, y no tener reglas no funciona en modo alguno.

Pero esta no puede ser la conclusión de este asunto —a menos que realmente entendamos como «adoración» una fidelidad racional—, debido a que *tener* reglas o *apoyar* o *aceptar* reglas no es, en absoluto, una solución de diseño. Tener reglas, tener toda la información e, incluso, tener buenas intenciones no basta para garantizar la acción correcta; el agente debe encontrar toda la materia correcta y utilizarla, incluso afrontando los adversos retos racionales diseñados para perturbar sus convicciones.

Tener y reconocer la fuerza de las reglas no es suficiente y, a veces, el agente se encuentra en mejores condiciones con menos. Douglas Hofstadter llamó la atención sobre un fenómeno que él denomina la «duda reverberante», la cual, en las discusiones teóricas más idealizadas, se estipula previamente como no utilizable. En lo que Hofstadter llama el «dilema del lobo», un obvio no dilema se convierte en un grave dilema sólo gracias al paso del tiempo y a la posibilidad de la «duda reverberante».

Imaginemos que veinte personas han sido seleccionadas en la clase de la Universidad, tú entre ellas. No sabes quiénes más han sido seleccionados… Lo único que sabes es que todos están conectados a un ordenador central. Cada uno de ellos se encuentra en un pequeño cubículo sentado en una silla, y frente a un botón en una pared que, por lo demás, está en blanco. Te dan diez minutos para decidir si pulsar o no el botón. Al terminar este período de tiempo se enciende una luz durante diez segundos y mientras la luz está encendida, puedes pulsar el botón o no pulsarlo. Todas las respuestas irán al ordenador central y un minuto después dará los resultados. Por fortuna, las consecuencias solamente pueden ser buenas. Si has pulsado el botón conseguirás cien dólares, sin condiciones. Si *nadie* pulsa el botón, entonces *todos* ganarán mil dólares. Pero si uno solo pulsa su botón, los que se han abstenido de pulsarlo no ganarán nada (Hofstadter 1985:752-753).

Obviamente no has pulsado el botón. ¿Correcto? Pero ¿qué pasaría si justamente una sola persona fuera demasiado precavida o dubitativa y comenzara a preguntarse si, después de todo, no era obvio que alguien lo haría? Todo el mundo debe aceptar que esta es una oportunidad extrema y todo el mundo debe reconocer que todos deben aceptar esto. Como Hofstadter señala es una situación «en la cual el más leve parpadeo de una duda ha llegado a simplificarse hasta convertirse en la más grave avalancha de dudas… Y una de las cosas molestas de este asunto es que mientras más brillante seas, más rápidamente y con más claridad, verás lo que hay que temer. Un grupo de amigos perezosos puede ser más propicio a abstenerse unánimemente y conseguir, de este modo, mayor beneficio que un grupo de lógicos de mente aguzada, todos pensando de manera perversa, recurrente y reverberante» (1985:753)**[179]**.

Enfrentados con un mundo donde tales situaciones conflictivas no son desconocidas, podemos reconocer lo atractivo que resulta un poco de religión

al viejo estilo, algo de un incuestionable dogmatismo, que convierte a los agentes en impenetrables a las sutiles invasiones de la hiperracionalidad. Crear algo parecido a una «posición disposicional» es sin duda uno de los objetivos del *Manual de primeros auxilios morales*, el cual, aunque lo imaginamos con el formato de un *consejo* para una audiencia atenta y racional, también puede ser visto como un método que no consigue su objetivo a menos que tenga el efecto de cambiar el «sistema operativo», no simplemente los «datos» (los contenidos de creencia o aceptación), de los agentes racionales a los cuales va dirigido. Por esto, para tener éxito en esa específica tarea debe dirigirse a las audiencias que constituyen sus objetivos con una gran precisión.

Puede haber, entonces, diferentes modelos del *Manual de primeros auxilios morales*, cada uno efectivo para un diferente tipo de audiencia. Esto abre una desagradable perspectiva ante los filósofos, por dos razones. En primer lugar, esto sugiere, contrariamente a sus austeros gustos académicos, que hay razón para prestar más atención a la retórica y otros medios de persuasión racional parcial o totalmente impuros; la *audiencia* racional ideal, a quien los que se dedican a la ética pueden presumir de dirigir sus reflexiones es, con todo, otra idealización dudosamente fructífera. Y, aún más importante, esto sugiere que lo que Bernard Williams llama el ideal de «transparencia» de la sociedad, esto es, «que el funcionamiento de sus instituciones éticas no debe depender de miembros de la comunidad que no entiendan correctamente cómo funcionan» (1985:101), es un ideal que puede ser políticamente inaccesible para nosotros. Aunque nos apartemos de una elitista construcción de mitos, y de doctrinas tan sistemáticamente carentes de ingeniosidad como el punto de vista que Williams llama «utilitarismo del gobierno» (1985:101), podemos encontrar —esto es una abierta posibilidad empírica, después de todo— que seremos extremadamente afortunados si hallamos una ruta racional y transparente desde lo que somos ahora, hasta lo que nos gustaría ser. El paisaje es abrupto y puede que no sea posible alcanzar los picos más altos desde la posición en la que nos encontramos hoy.

Repensar el diseño *práctico* de un agente moral, mediante el proceso de escribir varias versiones del *Manual de primeros auxilios morales*, puede, sin embargo, permitirnos dar sentido a algunos de los fenómenos en los que se basan las teorías éticas tradicionales. En primer término, podemos empezar a entender nuestra actual posición moral; y me refiero a la vuestra y la mía, en este mismísimo momento. Aquí están mis lectores, dedicando varias horas a leer mi libro (y yo estoy, sin duda, haciendo algo similar). ¿Debemos intentar

hacer una colecta de dinero para Oxfam, como institución humanitaria, o pasear en piquetes ante el Pentágono, o escribir cartas a nuestros senadores y representantes, acerca de varios asuntos? ¿Ha decidido conscientemente, tras unos cálculos, que había llegado el momento para un breve sabático apartado de los compromisos del mundo real, un período de retiro para leer un poco? ¿O su proceso de decisión —si no es un término muy rimbombante— consiste más bien en no obstaculizar algunos principios actuales «por omisión» que virtualmente le aseguran que sólo prestará atención a las interrupciones potenciales más excitantes de su vida personal, entre las cuales, me alegro de decirlo, se encuentran períodos dedicados a leer libros bastante difíciles?

Si es así, ¿es esta en sí misma una característica lamentable o algo sin lo cual nosotros, seres finitos, no podríamos pasar? Consideremos un «banco de pruebas» tradicional que la mayor parte de los sistemas de ética pasan con aplomo: resolver el problema de lo que usted debe hacer si se encuentra paseando, pensando en sus asuntos, y oye el grito de auxilio de un hombre que se está ahogando. Este es el problema fácil, una decisión local, convenientemente delimitada y bien *enmarcada*. El problema difícil es: ¿cómo voy hasta allí desde aquí? ¿Cómo puedo encontrar *justificadamente* una ruta desde nuestra difícil situación real hasta esta situación relativamente feliz y decidible sin dificultad? Nuestro problema prioritario, según parece, es que cada día, mientras tratamos desesperadamente de pensar en nuestros propios asuntos, escuchamos miles de gritos de ayuda, completados con volúmenes de información acerca de cómo podemos actuar para ayudar. ¿Cómo puede alguien sobre la Tierra establecer prioridades en medio de esta cacofonía? No precisamente a través de un proceso sistemático en el que se consideren todos los factores, sopesando las utilidades esperadas e intentando maximizarlas. No mediante cualquier generación y comprobación sistemática de máximas kantianas, demasiadas para que puedan ser consideradas.

No obstante, acabamos llegando hasta allí desde aquí. Pocos de nosotros nos quedamos paralizados por tales indecisiones durante largos períodos de tiempo. En conjunto, debemos resolver este problema de decisión permitiendo que una serie de opciones por omisión absolutamente «indefendible» proteja nuestra atención de todo, menos de nuestros actuales proyectos. La interrupción de aquellas opciones por omisión sólo puede ocurrir mediante un proceso dirigido por una heurística desordenada con moderadores arbitrarios y no reconocidos, que abusan de su autoridad.

Este ámbito de competición alienta naturalmente la intensificación del proceso. Con nuestra capacidad de atención estrictamente limitada, el problema con el que se enfrentan aquellos que desean que nos ocupemos de su tema favorito es esencialmente un problema de publicidad, de atraer la atención del bien intencionado. Esta competición entre memes es el mismo problema si lo vemos en el ámbito de la competición a gran escala de la política o en el primer plano de una deliberación personal. El papel de las fórmulas tradicionales para la discusión ética como directoras de la atención, o conformadoras de hábitos de imaginación moral, como metamemes *par excellence*, es un tema que merece un análisis ulterior.

CAPÍTULO 18
El futuro de una idea

I. *Elogio de la biodiversidad*

> Dios está en los detalles.
>
> *Ludwig Mies Van Der Rohe*

¿Cuánto tiempo tardó Johann Sebastian Bach en crear la *Pasión según San Mateo*? Una primera versión se estrenó entre 1727 y 1729, aunque la partitura que escuchamos hoy está fechada diez años más tarde e incorpora muchas revisiones. ¿Cuánto tiempo se tardó en crear a Johann Sebastian Bach? Había vivido cuarenta y dos años cuando tuvo lugar la primera audición de su primitiva versión y más de medio siglo cuando la última versión estuvo terminada. ¿Cuánto tiempo se tardó en crear la cristiandad, sin la cual la *Pasión según San Mateo* sería inconcebible, en el sentido literal del término, por Bach o por cualquiera otro? Aproximadamente dos milenios. ¿Cuánto tiempo ha sido necesario para crear el contexto social y cultural en el cual pudo nacer la cristiandad? Aproximadamente entre cien milenios y tres millones de años, según la fecha en la que fijemos el nacimiento de la cultura humana. ¿Y cuánto tiempo ha sido necesario para crear el *Homo sapiens*? Entre tres y cuatro billones de años, aproximadamente el mismo lapso de tiempo que fue necesario para que se crearan las margaritas y las percas, las ballenas azules y los búhos moteados. Billones de años de un *irreemplazable* trabajo de diseño.

Intuimos correctamente que existe un parentesco entre las más hermosas producciones del arte y de la ciencia, y las glorias de la bioesfera. William Paley estaba en lo cierto en lo siguiente: nuestra necesidad de encontrar una explicación acerca de cómo puede ser que el universo contenga tantas cosas maravillosamente diseñadas. La peligrosa idea de Darwin dice que *todas* existen como frutos de un árbol singular, el árbol de la vida, y los procesos que han producido todas y cada una de esas cosas, son todos, en el fondo, el mismo proceso. El genio desplegado por la madre naturaleza puede

desmontarse en muchos actos de microgenialidad, miópicos o ciegos, sin propósito, pero capaces, aunque en grado mínimo, del reconocimiento de una cosa buena (o mejor). Del mismo modo, el genio de Bach puede desmontarse en muchos actos de microgenialidad, pequeñas transiciones mecánicas entre estados cerebrales, que generan y que prueban, que descartan y revisan, y que vuelven a probar de nuevo. Entonces, ¿es el cerebro de Bach algo así como la proverbial imagen de los monos sentados ante la máquina de escribir? De ninguna manera, porque el cerebro de Bach en lugar de generar un inmenso número de alternativas, generó tan sólo un subgrupo evanescentemente pequeño de todas las posibilidades. Su genio puede ser medido, si se desea medir el genio, por la excelencia de este concreto subgrupo de candidatos generados. ¿Cómo fue capaz de moverse tan rápidamente el cerebro de Bach por el espacio de diseño, sin prestar atención apenas a las inmensas regiones vecinas llenas de diseños en los que no cabía depositar esperanza alguna? (Si desea explorar este territorio, siéntese al piano y trate, durante media hora, de componer una nueva melodía que valga la pena). El cerebro de Bach estaba exquisitamente diseñado a la manera de un programa heurístico para componer música, y el crédito de este diseño debe ser compartido; Bach tuvo fortuna con sus genes (procedía de una famosa familia de músicos) y también tuvo la fortuna de nacer en un medio cultural que había colmado su cerebro con los memes musicales de su época. Y, sin duda, Bach fue afortunado en muchos otros momentos de su vida al ser beneficiario de una u otra convergencia valiosa, hallada casualmente, sin ser expresamente buscada. Toda esta masiva contingencia produjo en Bach un vehículo único para recorrer y explorar, a velocidad de crucero, una parte del espacio de diseño que ningún otro vehículo podía explorar. No importa cuántos siglos o milenios de exploración musical nos queden por delante, nunca tendremos éxito en trazar caminos que dejen algo más que una huella en las inmensas extensiones del espacio de diseño. Bach es algo precioso no porque tuviera dentro de su cerebro una perla mágica de la materia de los genios, un gancho celeste, sino porque era, o contenía, una estructura absolutamente idiosincrásica de grúas, hecha a su vez de grúas, grúas y más grúas.

Al igual que Bach, la creación del resto del árbol de la vida difiere de los monos sentados ante la máquina de escribir por el hecho de haber explorado tan sólo un evanescente subgrupo de inmensas posibilidades. La eficiencia de la exploración ha sido creada una y otra vez, y son las grúas las que han acelerado la elevación a lo largo de los tiempos. Nuestra tecnología nos permite ahora acelerar nuestras exploraciones en cualquier parte del espacio

de diseño (no precisamente ingeniería genética, sino diseño asistido por ordenador de cualquier cosa imaginable incluyendo, por ejemplo, este libro, el cual nunca habría escrito sin el procesador de textos y el correo electrónico), pero nunca escaparemos de nuestra finitud o, dicho con mayor precisión, de nuestra trabazón con la realidad. La Biblioteca de Babel es finita pero inmensa, y nunca exploraremos todas sus maravillas y en cada punto debemos construir a modo de grúas, sobre las bases de lo que hemos edificado hasta la fecha.

Alertados por el omnipresente riesgo del reduccionismo insaciable, podemos considerar que parte de lo que valoramos es explicable por su cualidad de ser cosas diseñadas. Una pequeña intuición puede ayudarnos: ¿qué es peor, destruir el proyecto de alguien, aunque sea un modelo de la Torre Eiffel construido con miles de palillos de caramelos, o destruir su suministro de palillos de caramelos? Todo depende del objetivo del proyecto: si la persona disfruta diseñando y rediseñando, construyendo y reconstruyendo, entonces destruir su suministro de palillos de caramelos es peor; si no es el caso, entonces destruir el producto de diseño tan difícilmente conseguido será peor. ¿Por qué es bastante peor matar un cóndor que matar una vaca? (Traigo a colación este ejemplo, porque, sin tener en cuenta lo que otros piensen sobre lo que hay de malo en matar una vaca, estaremos de acuerdo en que es bastante peor matar a un cóndor, ya que la pérdida de nuestras actuales reservas de diseño en cóndores sería bastante mayor si los cóndores llegaran a extinguirse). ¿Por qué es peor matar a una vaca que matar a una almeja? ¿Por qué es peor matar a un árbol secoya que matar a una cantidad igual (en masa) de algas? ¿Por qué nos apresuramos a hacer copias de alta fidelidad de películas, discos, bandas sonoras y libros? *La Última cena* de Leonardo da Vinci se está desintegrando tristemente sobre una pared en Milán, a pesar de (a veces debido a) los esfuerzos por conservarla realizados durante siglos. ¿Por qué sería tan malo —puede ser peor— destruir todas las viejas fotografías que registran el aspecto que tenía hace treinta años como destruir hoy alguna porción del «original»?

Estas preguntas no tienen respuestas obvias e incontrovertibles, de modo que la perspectiva del espacio de diseño no explica todas las cosas acerca del valor, pero al menos nos permite comprender lo que sucede cuando tratamos de unificar, en una única perspectiva, nuestro sentido del valor. Por otro lado, nos ayuda a explicar nuestra intuición de que la unicidad o individualidad es «intrínsecamente» valiosa. Además, nos permite confirmar todas las inconmensurabilidades acerca de las cuales la gente habla. ¿Qué vale más,

una vida humana o la *Mona Lisa*? Hay muchos que darían su vida por salvar a este cuadro de la destrucción y muchos que *sacrificarían la vida de alguien*, si los colocaron ante esta alternativa. (¿Están armados los guardas del Museo del Louvre? ¿Qué pasos deben dar si fuera necesario?). ¿Merece la salvación del búho moteado la limitación de oportunidades en miles de vidas humanas afectadas? (Una vez más, los efectos retrospectivos se ven borrosos en la lejanía: si alguien ha invertido las posibilidades de su vida en llegar a ser un leñador y ahora nosotros eliminamos esa oportunidad, hemos devaluado su inversión de la noche a la mañana, tan seguramente —de hecho más seguramente— como si hubiéramos convertido sus ahorros de toda la vida en bonos basura, sin valor).

¿En qué «punto» comienza o termina una vida humana? La perspectiva darwiniana nos permite ver, con inequívoca claridad, por qué no hay en absoluto esperanza de *descubrir* una señal de cuentos de hadas, un salto en el proceso de la vida que pueda ser «tenido en cuenta». Necesitamos trazar líneas; necesitamos definiciones de la vida y la muerte debido a muy importantes propósitos morales. Las capas de dogma que, como una perla, van creciendo como una defensa alrededor de estos intentos fundamentalmente arbitrarios nos son familiares y se encuentran en un interminable proceso de reparación. Debemos abandonar la fantasía de que la ciencia o la religión puedan descubrir algunos hechos, hasta ahora bien ocultos, que nos digan exactamente dónde trazar estas líneas. No hay una vía «natural» para marcar el nacimiento de un «alma» humana, del mismo modo que no hay vía «natural» para marcar el nacimiento de una especie. Y, contrariamente a lo que en muchas tradiciones se insiste, creo que todos compartimos la intuición de que hay gradaciones de valor en el final de las vidas humanas. La mayor parte de los embriones humanos terminan, en aborto espontáneo; afortunadamente ya que la mayoría de ellos son *terata*, monstruos sin esperanza cuyas vidas serían imposibles. ¿Es éste un terrible mal? ¿Son las madres cuyos cuerpos abortan estos embriones culpables de crimen involuntario? Desde luego que no. ¿Qué es peor, tomar medidas «heroicas» para mantener vivos a una serie de niños gravemente deformados, o dar el igualmente «heroico» (aunque no alabado) paso de procurar que este niño muera tan rápidamente como sea posible y sin dolor? Yo no sugiero que el pensamiento darwiniano responda a todas estas cuestiones; sugiero que el pensamiento darwiniano nos ayuda a comprender por qué la tradicional esperanza de resolver estos problemas (algo así como encontrar un algoritmo moral) no tiene futuro. Debemos abandonar los mitos que hacen que estas

obsoletas soluciones parezcan inevitables. En otras palabras, necesitamos madurar.

Entre los preciosos artefactos que merecen ser preservados están todas las culturas. Existen aún varios miles de distintos lenguajes hablados diariamente en nuestro planeta, pero este número va disminuyendo rápidamente (Diamond 1992, Hale y otros 1992). Cuando un lenguaje desaparece es el mismo tipo de pérdida que la extinción de una especie, y cuando la cultura que era transmitida por ese lenguaje muere, ésta es una pérdida aún mayor. Pero aquí nos enfrentamos, de nuevo, con inconmensurabilidades y las respuestas no son fáciles.

Comencé este libro con una canción para mí muy querida que espero sobreviva para siempre. Espero que mi nieto la aprenda y la transmita a su nieto, aunque, al mismo tiempo, no creo y realmente no deseo que mi nieto crea las doctrinas que tan emotivamente se exponen en esa canción. Son demasiado simples. Son, en una palabra, erróneas, tan erróneas como las doctrinas de los antiguos griegos sobre los dioses y las diosas del Monte Olimpo. ¿Cree usted, literalmente, en un Dios antropomórfico? Si no cree, entonces debe coincidir conmigo en que la canción es una bella y reconfortante falsedad. No obstante, ¿es esta simple canción un valioso meme? Creo que, ciertamente, lo es. Es un modesto pero bello fragmento de nuestra herencia, un tesoro que debe ser conservado. Aunque debemos enfrentarnos con el hecho de que, así como hubo tiempos en que los tigres no hubieran sido viables, vendrán tiempos en que estos animales no serán viables excepto en los parques zoológicos y en otras reservas, y lo mismo puede decirse de muchos de los tesoros de nuestro patrimonio cultural.

El idioma galés se mantiene vivo por medios artificiales, igual que los cóndores. No podemos conservar *todos* los hechos y tesoros que han florecido en nuestro mundo cultural. No deberíamos desearlo. Fueron necesarios sistemas políticos y sociales represivos, repletos de muchas maldades, para crear el caldo de cultivo en el cual muchas de nuestras más grandes obras de arte pudieron florecer: esclavitud y despotismo (el pensamiento «ilustrado» a veces pudo haberlo sido), obscenas diferencias en los niveles de vidas entre los ricos y los pobres y enormes cantidades de ignorancia. La ignorancia es una condición necesaria para muchas cosas excelentes. La alegría infantil de ver lo que Santa Claus ha traído por Navidad es una especie de alegría que debe extinguirse pronto en cada niño con la pérdida de la ignorancia. Cuando esta niña crece, puede transmitir esta alegría a sus propios hijos, pero también debe reconocer a tiempo cuándo ha perdido su valor.

El punto de vista que expreso aquí tiene claros antecesores. El filósofo George Santayana era un católico ateo, si tal cosa puede ser posible. Según Bertrand Russell (1945:811), William James denunció, en cierta ocasión, las ideas de Santayana como la «perfección de la decadencia», y se puede comprender la razón por la que algunos llegan a sentirse ofendidos por esta rama del esteticismo: una profunda apreciación de todas las fórmulas, ceremonias y adornos de su herencia religiosa, pero sin fe. La posición de Santayana fue apropiadamente caricaturizada: «No hay Dios y María es su Madre». Pero ¿cuántos de nosotros estamos atrapados en un dilema semejante, amando nuestra herencia, firmemente convencidos de su valor, pero incapaces de mantener la convicción de que todo esto es verdad? Nos enfrentamos a una difícil elección. Debido a que valoramos esta herencia, estamos dispuestos a conservarla, en un estado precario o «desnaturalizado», en iglesias, catedrales y sinagogas, construidas para acoger a grandes concentraciones de fieles, y ahora en vías de convertirse en museos culturales. No hay realmente demasiada diferencia entre lo que hacen los guardias reales montando su pintoresca guardia en la Torre de Londres y los cardenales cuando caminan con sus magníficos ropajes y se reúnen para elegir al nuevo Papa. En ambos casos están manteniendo vivas tradiciones, rituales, liturgias, símbolos, que de otro modo desaparecerían.

Pero, en todos estos credos, ¿no ha habido un tremendo renacer de la fe fundamentalista? Sí, desafortunadamente lo ha habido y pienso que no hay en el planeta fuerzas más peligrosas para nosotros que el fanatismo de los fundamentalistas, sean de la especie que sean: protestantismo, catolicismo, judaísmo, islamismo, hinduismo y budismo, así como una incontable serie de infecciones más pequeñas. ¿Existe en este asunto un conflicto entre la religión y la ciencia? Ciertamente lo hay.

La peligrosa idea de Darwin ayuda a crear una condición en la memosfera que, a la larga, amenaza ser tan tóxica para estos memes como la civilización, en general, lo ha sido para una gran cantidad de mamíferos salvajes. ¡Salvemos los elefantes! Sí, naturalmente, pero no *por todos los medios*. No forzando, por ejemplo, a los pueblos de África a vivir como en el siglo XIX. Esta no es una comparación ociosa. La creación de grandes reservas para animales en África ha provocado a menudo la dislocación —y, en último término, la destrucción— de poblaciones humanas. (Una horripilante descripción de estos efectos colaterales se puede encontrar en Colin Turnbull 1972, sobre la suerte del pueblo Ik en África). Aquellos que piensen que podemos conservar, a toda costa, los prístinos medios ambientes de los

elefantes deben tener en cuenta los costes que supondría hacer retroceder a los Estados Unidos a las condiciones primitivas en las que los búfalos vagaban libremente y los ciervos y los antílopes retozaban. Debemos encontrar un término medio.

Me encanta la versión de la Biblia del rey Jaime. Mi propio espíritu se aparta con temor de un Dios que es El o Ella, del mismo modo que mi corazón se acongoja cuando veo un león paseando neurótico, arriba y abajo, en la pequeña jaula de un zoo. Ya sé muy bien que el león es bello pero peligroso; si le dejo vagar libremente me mataría; la seguridad exige que se le mantenga en una jaula. La seguridad exige que las religiones sean colocadas también en jaulas, cuando sea absolutamente necesario. No podemos vernos forzados a aceptar la circuncisión de la mujer, la situación de segundo orden de la mujer en el catolicismo romano y en el mormonismo, por no decir nada de su situación en el islam.

La reciente toma de posición del Tribunal Supremo de Estados Unidos al declarar inconstitucional la ley de Florida que prohíbe el sacrificio de animales en los rituales de la santería (una religión afrocaribeña que incorpora elementos de la tradición yoruba y del catolicismo romano) es un caso límite, al menos para muchos de nosotros. Estos rituales nos resultan ofensivos a muchos de nosotros, aunque el manto protector de las tradiciones religiosas les asegura nuestra tolerancia. Es inteligente respetar estas tradiciones. Después de todo, forman parte de la biosfera.

¡Salvemos a los bautistas! De acuerdo, pero no *por todos los medios*. No si estos medios toleran la errónea información sobre el mundo natural que se da deliberadamente a los niños. Según una reciente encuesta, el 48 por ciento de la gente en Estados Unidos cree hoy que el libro del Génesis es literalmente verdad. Y el 70 por ciento cree que la «ciencia de la creación» debe enseñarse en las escuelas junto con la enseñanza de la evolución. Algunos escritores recomiendan una política que permitiría a los padres «rechazar» los materiales que no desean que les sean enseñados a sus hijos. ¿Debe ser enseñada en las escuelas la evolución? ¿Debe ser enseñada la historia? Informar erróneamente a un niño es un delito terrible.

Una fe, como las especies, puede evolucionar o extinguirse cuando el medio ambiente cambia. En cualquier caso, no es un proceso suave. Observamos en cada una de las subespecies cristianas la batalla de los memes: ¿deben ser ordenadas las mujeres? ¿Debemos volver a la liturgia en latín?; lo mismo puede observarse en las variedades del judaismo y del islam. Debemos tener una mezcla similar de respeto y de precaución autoprotectora

con respecto a los memes. Es ya una práctica aceptada, aunque tendemos a mirar hacia otro lado, en lo que se refiere a sus implicaciones. Predicamos libertad religiosa, pero sólo hasta cierto punto. Si nuestra religión aboga por la esclavitud o la mutilación de las mujeres o el infanticidio o pone un precio a la cabeza de Salman Rushdie porque se siente insultada, entonces nuestra religión tiene una característica que no puede ser respetada. Significa un riesgo para todos.

Es bonito tener osos grises y lobos en el bosque. Ya no son una amenaza; podemos coexistir pacíficamente si sabemos cómo hacerlo. La misma conducta puede aplicarse a nuestra tolerancia política y a la libertad religiosa. Somos libres de conservar o crear cualquier credo religioso que deseemos, siempre que no represente una amenaza pública. Estamos juntos en la Tierra y hemos de aprender a acomodarnos. Los memes hutteritas son lo suficientemente «listos» como para no incluir ningún meme acerca de la virtud de destruir a los forasteros. De haberlo hecho, los hubiéramos combatido. Toleramos a los hutteritas porque se hacen daño solamente a sí mismos, aunque podamos insistir en que tenemos el derecho de imponer alguna apertura en la escolarización de sus hijos. Otros memes religiosos no son tan benignos. El mensaje es claro: a aquellos que no se acomoden, que no se atemperen, que insistan en mantener viva únicamente la más pura e indomable casta de su herencia, nos veremos obligados, a regañadientes, a encerrarlos o desarmarlos y haremos todo lo posible por inhabilitar los memes que utilizan para luchar. La esclavitud está fuera de los límites de lo permisible. El abuso de los niños está fuera de los límites. La discriminación está fuera de los límites. Las sentencias de muerte a aquellos que blasfeman contra una religión (complementada con beneficios o recompensas para aquellos que la ejecuten) están fuera de los límites. No es civilizado y no merece más respeto en el nombre de la libertad religiosa que cualquier otra incitación al crimen a sangre fría[180].

A aquellos de nosotros que vivimos vidas plenas, e incluso emocionantes, difícilmente debería sorprendernos observar que la gente en un mundo subdesarrollado, con ínfimas condiciones sociales —incluidos los más sórdidos rincones de nuestro propio mundo—, se inclina hacia el fanatismo de un tipo o de otro. ¿Aceptaríamos dócilmente una vida insignificante de pobreza, sabiendo lo que sabemos hoy acerca del mundo? La tecnología de la infoesfera ha hecho concebible recientemente que todos los que habitan el globo sepan aproximadamente lo que nosotros sabemos (con bastante distorsión). Mientras no podamos facilitar a todo el mundo un medio

ambiente en el cual el fanatismo no tenga sentido, podemos esperar más y más de esta plaga. Pero no tenemos que aceptarlo ni tampoco respetarlo. Tomando unos cuantos sorbos de la medicina darwiniana (Williams y Nesse 1991) podemos dar los pasos necesarios para conservar lo que es valioso en cada cultura sin mantener vivas (o virulentas) todas sus debilidades.

Cabe apreciar la belicosidad de los espartanos sin aspirar a reintroducir su sistema; podemos maravillarnos ante todos los sistemas de atrocidades introducidos por los mayas sin que, en ningún momento, lamentemos la desaparición de aquellas prácticas. La salvación para la posteridad de estos desusados artefactos culturales debe ser académica pero no en reservas humanas de caza: estados étnicos o religiosos bajo una dictadura. El griego antiguo y el latín no son, desde hace mucho tiempo, lenguas vivas, pero desde el punto de vista académico se han conservado el arte y la literatura de la antigua Grecia y de Roma. Petrarca, en el siglo XIV, alardeaba de los volúmenes de filosofía griega que tenía en su biblioteca personal; no pudo leerlos entonces porque el conocimiento de la antigua Grecia casi había desaparecido totalmente en el mundo que le tocó vivir, pero conocía su valor y se esforzó por restablecer los conocimientos que revelaban sus secretos.

Mucho tiempo de que existiera ciencia, e incluso filosofía, había religiones. Han servido a muchos propósitos (sería un error de reduccionismo insaciable ver un solo propósito, un *summum bonum* singular al cual directa o indirectamente todas habrían servido). Han inspirado a mucha gente para seguir unas vidas que han incrementado, de manera inconmensurable, las maravillas de nuestro mundo, y han inspirado a mucha más gente a llevar unas vidas que estuvieron, dadas sus circunstancias, más llenas de significado, que fueron menos dolorosas de lo que podían haber sido. La pintura de Brueghel *La caída de Icaro* muestra, en primer plano, a un labrador con su arado y a un caballo en la ladera de una colina, mientras que un bello barco de vela se desliza al fondo del cuadro y dos piernas blancas, que casi pasan inadvertidas, desaparecen en el mar, con un discreto chapoteo. Este cuadro inspiró a W. H. Auden uno de mis poemas favoritos.

«MUSÉE DES BEAUX ARTS»

Jamás se equivocaban acerca del sufrimiento,
los Viejos Maestros: cómo comprendían
su posición humana; cómo tiene lugar
mientras algún otro está comiendo o abriendo una ventana o
sencillamente andando aburridamente;
cómo, mientras los ancianos están esperando reverente, apasionadamente
el milagroso nacimiento, siempre tiene que haber
niños que no tenían ningún deseo especial de que se produjera patinando

sobre un estanque en el borde del mundo:

jamás olvidaron
que incluso el temible martirio ha de llegar a su fin
de cualquier manera en una esquina, en algún punto deseado
donde los perros viven su perruna vida y el caballo del torturador
se rasca su inocente trasero contra un árbol.

En el *Icaro* de Brueghel, por ejemplo: como se aleja todo
calmadamente del desastre; el hombre del arado puede
que haya oído el chapoteo, el grito desesperado,
pero para él no era un fracaso importante; el sol brillaba
como debía sobre las blancas piernas que desaparecían en la verde
agua; y el valioso y delicado barco que tenía que haber visto
algo asombroso, un muchacho cayendo del cielo,
tenía que llegar a alguna parte y seguía calmoso su camino.

Este es nuestro mundo, donde el sufrimiento nos importa, si es que importa algo. Las religiones han proporcionado la comodidad de la pertenencia y la camaradería a muchos que, de otro modo, habrían pasado su vida solos, sin gloria y sin aventura. En el mejor de los casos, las religiones han llamado la atención sobre el amor, y lo han hecho realidad para gente que, de otra forma, no lo habría conocido, ennobleciendo las actitudes y reavivando los espíritus de los perseguidos en el mundo. Otra cosa que las religiones han logrado, sin ser su *raison d'être*, ha sido mantener al *Homo sapiens* bastante civilizado, durante largo tiempo, enseñándole a reflexionar más sistemáticamente y con mayor precisión sobre nuestra posición en el mundo. Hay mucho más que aprender. Hay ciertamente un tesoro de verdades mal apreciadas, embebidas en las culturas en peligro en el mundo moderno, diseños que han acumulado detalles a lo largo de los tiempos de historia idiosincrásica, y debemos dar los pasos necesarios para recordarlas y estudiarlas antes de que desaparezcan ya que, como el genoma de los dinosaurios, una vez desaparecido, será virtualmente imposible de recuperar.

No debemos esperar que este tipo de respeto sea satisfactorio para aquellos que de todo corazón encarnan los memes que nosotros honramos con nuestra atenta —aunque no rendida— dedicación académica. Por el contrario, muchos de ellos verán todo lo que no sea una conversión entusiasta a sus propios puntos de vista como una amenaza, incluso como una intolerable amenaza. No debemos subestimar el sufrimiento que tal enfrentamiento causa. Contemplar y tener que participar en la contracción o evaporación de apreciadas características de nuestra herencia es un dolor que solamente nuestra especie puede experimentar y seguro que hay pocos dolores más terribles. Pero no disponemos de una alternativa razonable y a aquellos cuya

visión del mundo les dicta que no deben coexistir pacíficamente con los demás, tendremos que mantenerlos en cuarentena del mejor modo posible, minimizando el dolor y el daño, intentando siempre dejar abierto uno o dos caminos que puedan llegar a parecer aceptables.

Si deseáis enseñar a vuestros hijos que ellos son herramientas de Dios, lo mejor no es enseñarles que son rifles de Dios o, de lo contrario, me opondré firmemente: vuestra doctrina no tiene gloria, ni derechos especiales, ni mérito intrínseco e inalienable. Si estáis dispuestos a insistir en enseñar a los niños falsedades —que la Tierra es plana, que el «hombre» no es un producto de la evolución por selección natural— entonces debéis esperar, al menos, que aquellos de nosotros que disfrutamos de libertad de expresión hablemos sin trabas para describir vuestras enseñanzas como una siembra de falsedades, o que intentemos demostrárselo a vuestros hijos en la primera oportunidad. Nuestro futuro bienestar —el bienestar de todos en el planeta— depende de la educación de nuestros descendientes.

¿Qué queda entonces de todas las glorias de nuestras tradiciones religiosas? Ciertamente deben conservarse, igual que las lenguas, el arte, las costumbres, los rituales y los monumentos. Los parques zoológicos son ahora considerados, más o menos, como refugios de segunda clase para las especies en peligro, pero al menos son refugios y lo que conservan es irreemplazable. Lo mismo puede decirse de los complejos memes y sus expresiones fenotípicas. Muchas hermosas iglesias de Nueva Inglaterra, costosas de mantener, están en peligro de destrucción. ¿No sería aconsejable desconsagrar estas iglesias y convertirlas en museos o reciclarlas para cualquier otro uso? La segunda opción es, al menos, preferible antes que su destrucción. Muchas congregaciones se enfrentan a una cruel elección: cuesta mucho mantener su casa de adoración en todo su esplendor, por lo que sólo una pequeña parte de sus diezmos se entrega a los pobres. La Iglesia católica se ha enfrentado a estos problemas durante siglos y ha mantenido una posición que considero defendible, aunque no lo sea a primera vista: cuando gasta sus tesoros para dorar los candelabros, en lugar de aportar más alimentos y mejores refugios a los pobres de la parroquia, tienen una visión diferente de las cosas que merecen la pena en esta vida. Nuestro pueblo, argumentan, se beneficia más teniendo un lugar de esplendor donde adorar que teniendo un poco más de alimentos. Cualquier ateo o agnóstico que encuentre ridículo este análisis de coste-beneficio, puede considerar la opción de apoyar el trasvase de todos los fondos benéficos y gubernamentales dedicados a los museos, orquestas sinfónicas, bibliotecas y laboratorios científicos al esfuerzo por facilitar más

alimentos y mejores condiciones de vida a los menos favorecidos. Una vida humana que merezca ser vivida no es algo que pueda medirse incontrovertiblemente y en eso radica su gloria.

Y aquí radica el problema. ¿Qué sucedería, podríamos preguntarnos, si la religión fuera conservada en zoos culturales, en bibliotecas, en conciertos y en exhibiciones? Está sucediendo; los turistas se congregan para observar las danzas tribales de los indígenas americanos, y para los espectadores esto es folklore, una ceremonia religiosa, ciertamente, merecedora de respeto, pero también un ejemplo de un meme complejo al borde de la extinción, al menos de su fase fuerte y ambulatoria; un meme que se ha convertido en un inválido, que se mantiene vivo gracias a sus guardianes. ¿Nos ha dado la peligrosa idea de Darwin algo a cambio de las ideas que ha cuestionado?

En el capítulo 3, cité al físico Paul Davies, quien proclamaba que el poder de reflexión de las mentes humanas puede ser «no un detalle trivial ni un producto colateral menor de fuerzas insensatas y sin propósito», y sugería que el hecho de ser un producto colateral de fuerzas insensatas y sin propósito no descalificaba su importancia. Y yo he argumentado que Darwin nos ha mostrado cómo, de hecho, *todo* lo que tiene importancia es justamente un producto de este tipo. Spinoza llamó Dios o naturaleza (*Deus sive Natura*) a su ser más elevado, expresando una suerte de panteísmo. Ha habido muchas variedades de panteísmo, pero todas carecen de una convincente *explicación* acerca de cómo Dios se distribuye en la totalidad de la Naturaleza. Como vimos en el capítulo 7, Darwin nos ofrece una explicación: es en la distribución del diseño a través de la naturaleza, creando, en el árbol de la vida, una creación absolutamente única e irreemplazable, un patrón real en la inconmensurable extensión del espació del diseño que nunca puede ser exactamente duplicado en sus muchos detalles. ¿Qué es trabajo de diseño? Es esa admirable conjunción de azar y necesidad, que sucede en un trillón de lugares a la vez y a un trillón de diferentes niveles. ¿Qué milagro la ha causado? Ninguno. Sencillamente sucedió porque sí, en la plenitud de los tiempos. Podríamos incluso decir, en cierto modo, que el árbol de la vida se ha creado a sí mismo. No en una instántanea exhalación, sino lenta, muy lentamente, durante billones de años.

¿Es este árbol de la vida un Dios que merece adoración? ¿Merece que se le eleven plegarias? ¿Temor? Probablemente no. Pero este árbol hace que la hiedra se retuerza y que el cielo sea tan azul, por lo que quizá la canción que tanto me gusta dice una verdad, después de todo. El árbol de la vida ni es perfecto ni infinito en el espacio o el tiempo, pero es real, y si no es lo que

pensaba San Anselmo, «Un ser más grande que todo lo que uno pueda concebir», es seguramente un ser que es mayor que cualquier cosa que cualquiera de nosotros concebiría en un detalle merecedor de su detalle. ¿Es algo sagrado? Sí, afirmo con Nietzsche. Yo no puedo rezarle, pero puedo apoyar la afirmación de su magnificencia. Este mundo es sagrado.

2. *El ácido universal: manéjese con cuidado*

No se trata de negar, llegados a este punto, que la idea de Darwin es un disolvente universal capaz de penetrar hasta el corazón de cualquier cosa a la vista. La pregunta es: ¿qué deja tras de sí este ácido? He tratado de demostrar que una vez que el ácido pasa a través de todas las cosas, nos deja con unas versiones más sólidas y más fundamentadas de nuestras ideas más importantes. Algunos de los detalles tradicionales perecen, y algunos de éstos son pérdidas lamentables, pero en cuanto al resto ¡en buena hora nos libramos! Lo que permanece es más que suficiente para seguir añadiendo.

Durante las tumultuosas controversias que han acompañado la evolución de la peligrosa idea de Darwin siempre ha habido, en cada escenario, un desafío nacido del miedo: «¡*Nunca* logrará explicar *esto*!». Y el reto ha sido aceptado: ¡Mire cómo lo hago! Y a pesar de las enormes inversiones emocionales —y en parte debido a ellas— que los oponentes han hecho para ganar en ambos lados del argumento, el cuadro ha ido quedando cada vez más claro. Ahora conocemos *bastante* mejor el sentido de lo que es un algoritmo darwiniano, mejor de lo que Darwin hubiera soñado. La intrépida ingeniería revertida nos ha llevado a un punto desde donde podemos confiadamente valorar las posiciones rivales acerca de lo que exactamente sucedió en este planeta hace billones de años. Los «milagros» de la vida y la consciencia resultan ser mejores de lo que imaginábamos cuando estábamos seguros de que eran inexplicables.

Las ideas expresadas en este libro son solo el comienzo. Esta ha sido una introducción al pensamiento darwiniano, sacrificando detalles, una y otra vez, para aportar una mejor apreciación de la conformación global de la idea de Darwin. Pero, como ha dicho Mies van der Rohe, Dios está en los detalles. Recomiendo añadir precaución al entusiasmo que espero haber infundido en el lector. He aprendido de mi propia embarazosa experiencia cuán fácil es confeccionar explicaciones darwinianas notablemente convincentes que se desvanecen ante un examen más detallado. El verdadero aspecto peligroso de la idea de Darwin se deriva de su capacidad de seducción. Las versiones de

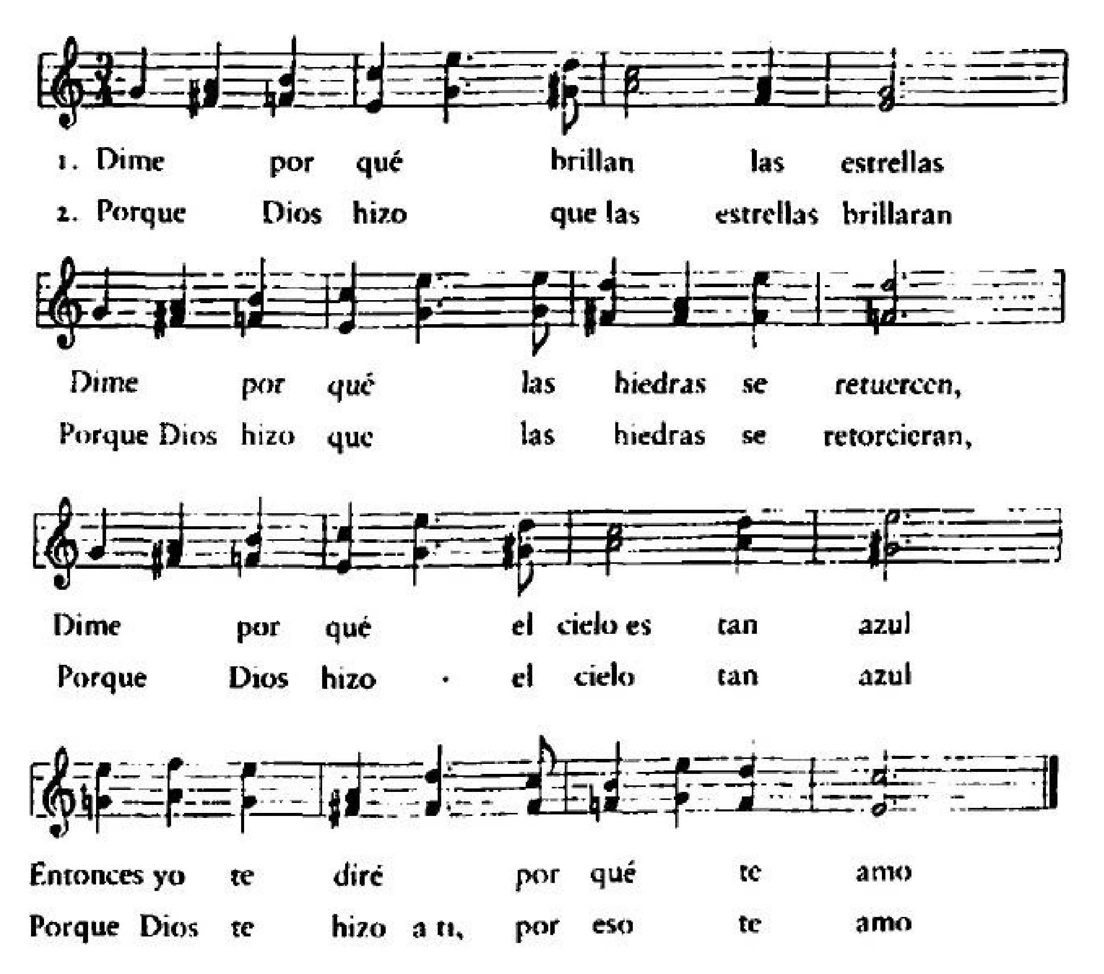
segunda categoría de las ideas fundamentales continúan hechizándonos, por lo que debemos estar muy atentos, corrigiendo lo que sea necesario. El único camino para evitar este error es aprender de los errores que hemos cometido.

Un meme que aparece bajo muchos disfraces en el folklore mundial es el cuento del amigo con aspecto terrorífico inicialmente confundido con un enemigo. «La Bella y la Bestia» es una de las especies mejor conocida de esta historia. Como contrapeso está la historia del «lobo disfrazado de oveja». Entonces, ¿qué meme deseamos utilizar para expresar nuestro juicio acerca del darwinismo. ¿Es verdaderamente un «lobo disfrazado de oveja»? En tal caso, ha de rechazarse y deberemos luchar contra él, más vigilantes que nunca ante la seducción de la idea de Darwin, que es verdaderamente peligrosa. O ¿resulta que la idea de Darwin, después de todo, es justamente lo que necesitamos para conservar y explicar los valores que amamos? He completado mi caso para la defensa: La Bestia es, de hecho, una amiga de la Belleza y de hecho bastante bella a su manera. ¡Lector, a ti te corresponde ahora ser el juez!

Apéndice

Dime por qué

Canción popular



(La línea armónica se canta generalmente con las voces más altas una octava por encima de la melodía.)

Glosario

ADAPTACIÓN: La capacidad de las especies para modificarse en el curso de su vida en respuesta a los cambios que acaecen en su medio ambiente. La adaptación permite mejorar su aptitud para la supervivencia, comparada con la que tenía originalmente su fenotipo.

ADAPTACIONISMO: Dentro de la teoría de la evolución, son *adaptacionistas* los que sostienen que la mayoría, si no todas, de las características orgánicas son adaptativas y han aparecido específicamente como resultado directo de la selección natural.

ALELO: Cada una de las variantes de un grupo de genes que puede ocupar un *locus* determinado en el cromosoma y que controlan el mismo carácter.

ALGORITMO: Un procedimiento completo, y nada ambiguo, para resolver un problema específico en un número finito de pasos.

ALGORITMO DETERMINSTA: Un algoritmo en el que, dado el mismo *input* de información, producirá siempre el mismo *output* de información, cuando se aplica correctamente.

ALGORITMO ESTOCÁSTICO: Dado el mismo *input* de información, no siempre producirá el mismo *output* de información, aunque se aplique correcta mente.

ALGORITMOS RANDOMIZADOS DE MICHAEL RABIN: Este profesor de ciencia de la computación ha desarrollado sofisticados algoritmos para proteger de intrusos los archivos de los ordenadores, especialmente mediante la aplicación de la aleatorización a los algoritmos de la computación.

ALGORITMO GENÉTICO: La utilización de técnicas evolucionistas para diversificar, combinar y seleccionar opciones con el objetivo de mejorar los resultados, siguiendo los métodos de la selección natural.

APTITUD PARA SOBREVIVIR: La habilidad de un organismo para sobrevivir en su medio ambiente, en relación con otras criaturas que allí se encuentran.

APTITUD INCLUSIVA PARA SOBREVIVIR: El éxito global de un individuo en la reproducción de sus genes, directamente a través de sus

crías o indirectamente ayudando a su parentela, con la que comparte genes y a la que «incluye» en su aptitud para sobrevivir.

ANAGÉNESIS: Evolución dentro de un linaje. Véase *Cladogénesis*.

ARMAMENTO, CARRERA DE: Proceso interactivo por el cual una especie que actúa como presa cambia en respuesta a modificaciones acaecidas en la otra especie, que actúa como predadora.

ATRACTOR: Un punto hacia el cual un sistema tiende a moverse y que funciona como un objetivo, sea deliberado o limitado por leyes o parámetros de dicho sistema.

BEHAVIORISMO: También conocida como *conductismo*, es una escuela de psicología que acepta la evidencia objetiva de la conducta —medida en función de las respuestas a estímulos— como el único objetivo de su investigación y el único fundamento de su teoría, sin referencia a la experiencia consciente.

BIOSFERA: Término que incluye a los seres vivos en su medio ambiente. Véase *Hipótesis de Gaia*.

CAOS: Un sistema cuya conducta a largo plazo es impredecible, de modo que mínimas modificaciones en la exactitud del valor inicial divergen hacia cualquier lugar de su *state space*.

CLADOGÉNESIS: Evolución que da origen a la bifurcación de un linaje. Véase *Anagenesis*.

CODON: La unidad del código genético que determina la síntesis de aminoácidos. Cada codon consiste en una sección de la molécula de ADN y el orden de los códones a lo largo de la molécula determina el orden de los aminoácidos en cada proteína sintetizada por la célula.

COMPLEJIDAD: La interacción de muchas partes que da origen a propiedades que no se encuentran en los agentes individuales.

CONDICIONAMIENTO OPERANTE O INSTRUMENTAL: El término «operante» se refiere a la noción de que los seres humanos aprenden operando sobre su medio ambiente. Nos comportamos, sacamos consecuencias de este comportamiento y las utilizamos para modular nuestra conducta futura. Una conducta que va seguida de una consecuencia positiva tenderá a ser repetida («Ley del efecto»).

CONDUCTISMO: Véase *Behaviorismo*.

CONSTRAINTS: El término se refiere a las 'limitaciones' que impone la naturaleza o el ser humano para impedir que tengan lugar ciertas acciones.

CONTINGENCIA: Una condición bajo la cual tiene lugar una conducta y sus consecuencias.

DEME: En la biología teórica, una entidad geográfica natural que evoluciona aisladamente.

EFECTO MARIPOSA: Define la posibilidad de que pueda producirse un gran cambio en un sistema a partir de una leve desviación de las condiciones iniciales. Según esta teoría el aleteo de una mariposa (*butterfly effect*) en la Amazonia podría conducir a cambios en la localización de un tifón.

EFECTO BALDWIN: La tendencia de los organismos que aprenden o adquieren características útiles para tener éxito, lo que conduce a una probabilidad más elevada de su reproducción y a una fijación de estas útiles características en la población, a pesar de la ausencia de una herencia directa de dichas características.

ENTROPÍA: Es una medida del grado de desorden de un sistema. En los procesos irreversibles, la entropía se va incrementado y la energía disponible se va disipando. La entropía total del universo aumenta continuamente con tendencia a alcanzar un máximo, el cual correspondería a un completo desorden de las partículas —átomos y moléculas—, si se asume que el universo es un sistema aislado. Véase *Muerte calórica del universo*.

EQUILIBRIO PUNTUADO: La evolución puntuada —en contraposición al *gradualismo*— ocurre cuando el incremento de cambio es muy grande comparado con el del tiempo en intervalos discretos, mientras que en la mayor parte del tiempo no habría prácticamente cambios (períodos de «estasis»).

ESPECIACIÓN: Proceso que da lugar a la formación de las especies.

ESPECIES EUSOCIALES: Especies en las que existe una división del trabajo reproductivo en la que los trabajadores no se reproducen a sí mismos, pero trabajan en representación de los individuos implicados en la reproducción.

EUCARIOTA: Organismo que posee un núcleo bien diferenciado, con más de un cromosoma, separado del citoplasma, con organelos (mitocrondrias, aparato de Golgi) y ribosomas.

EX-ADAPTACIÓN: Características de un organismo, en principio neutrales desde el punto de vista selectivo, y que son posteriormente adaptadas para alguna función de consecuencias selectivas.

FEEDBACK: En un sistema en el que tiene lugar una transformación hay *inputs* y *outputs*. Los *inputs* son el resultado de la influencia del medio ambiente sobre el sistema y los *outputs* son la influencia del sistema sobre

el medio ambiente. El *input* y el *output* están separados por una duración de tiempo, como antes y después, pasado y presente. El *feedback* se concibe como un bucle que, por fuera del sistema, transmite al *input* del sistema la información acerca del resultado (*output*) de la transformación. Si estos datos facilitan o aceleran la transformación en la misma dirección que siguieron los resultados anteriores, el *feedback* se considera positivo. De lo contrario, el *feedback* se considera negativo y se estabiliza el sistema.

FENOTIPO: La expresión detectable de la interacción del genotipo y su medio ambiente que constituyen los caracteres visibles de un organismo.

FRACTAL: Un sistema que posee detalles similares en todas sus escalas, dando lugar a configuraciones intrincadas y características inesperadas.

GEN: La unidad básica de la herencia mendeliana, que representa una región contigua de ADN (o ARN en algunos virus) correspondiente a una (con menos frecuencia, a dos o más) unidad de transcripción. La secuencia característica de los nucleótidos a lo largo de la molécula de ácido nucleico representa una unidad funcional de la herencia. Los genes pueden ser divididos en aquellos que codifican polipéptidos (genes estructurales), aquellos en los que se realiza la transcripción en ARN pero no son trasladados en proteínas, y aquellos en los que su significación funcional no exige que se produzca su transcripción.

GENOTIPO: La combinación de genes a partir de los cuales se construye un organismo.

GRADUALISMO: Teoría que sostiene que La evolución ocurre cuando el incremento de cambio es pequeño comparado con el incremento de tiempo. Véase *Equilibrio puntuado*.

HALTING PROBLEM: Véase *Turing, máquina de*.

HAMMING DISTANCE: La distancia entre dos «palabras codificadas», con referencia a datos digitalizados y codificados en *bits*, es el número de cambios necesarios en un *bit* para transformarlo en otro bit. La distancia entre 10100111 y 11000101 es 3. La distancia Hamming de un conjunto de «palabras codificadas» es la distancia más corta entre todos los posibles pares de palabras codificadas distintas en el conjunto.

MUERTE CALÓRICA DEL UNIVERSO: La segunda ley de la termodinámica predice el fin del universo —si se considera como un sistema cerrado— como consecuencia del agotamiento del calor transferible, al alcanzarse un estado de entropía máxima en el que todas

las cosas se encontrarían a la misma temperatura, sin energía disponible para ser utilizada.

HILL-CLIMBÍNG: La habilidad de una mutación para incrementar la capacidad de sobrevivir en su medio ambiente.

HIPERCICLOS: Un modelo de sistema macromolecular autorreproductivo, descrito por M. Eigen y P. Schuster [1979], con una determinada organización entre las matrices del ARN y las enzimas de la replication, que asegura su estabilidad y supervivencia durante la competición.

HIPERESPACIO: Un espacio con más de tres dimensiones. Cualquier construcción geométrica cuyos puntos se caracterizan por más de tres valores. Es difícil para la mente humana imaginarse operando en tales espacios.

HIPÓTESIS DE GAIA: Hipótesis propuesta a finales de la década de los sesenta y comienzos de los setenta por James Lovelock —científico británico dedicado a los estudios de la atmósfera, tanto terrestre como extraterreste— y Lynn Margulis, microbióloga norteamericana. A la pregunta de por qué la Tierra, en comparación con otros planetas, ha sido transformada en un sistema viviente que autoevoluciona y se autorregula. Lovelock y Margulis respondieron formulando la hipótesis de que la Tierra sería, en sí misma, un ser vivo. Toda la materia viva de la Tierra, desde las ballenas a los virus, desde los robles hasta las algas, podría ser considerada como constituyente de una entidad viva singular. El novelista William Golding, autor de *El señor de las moscas*, vecino de Lovelock, le propuso como título de su hipótesis el nombre de Gaia, la diosa griega que extrajo a nuestro mundo del caos.

HOMEOSTASIS: La capacidad de un sistema para autorregularse y mantener un estado concreto.

HOMOLOGÍAS: Similitudes en las secuencias del ADN entre individuos de la misma especie o de diferentes especies.

HOMÓLOGO: Un par de cromosomas que contiene la misma secuencia lineal de genes, cada una derivada de un progenitor.

IK: Tribu del norte de Uganda, conocida como «The Mountain People» (‘gente de las montañas’), que fue desplazada a la fuerza de su hábitat en las montañas boscosas y trasladada a un nuevo asentamiento y una nueva vida sedentaria como agricultores.

INGENIERÍA REVERTIDA: El proceso de desarmar un producto o artefacto con el fin de determinar cómo fue diseñado.

INTELIGENCIA ARTIFICIAL: La capacidad de un ordenador para realizar tareas habitualmente asociadas con procesos intelectuales elevados característicos de los seres humanos, como razonar, descubrir significados, generalizar o aprender a partir de experiencias pasadas.

KLUGE: Así se califica a algo no diseñado en su globalidad, sino constituido a partir de componentes disponibles. Término usado con frecuencia en ingeniería.

LEY DE DOLLO: Esta ley establece que la evolución es irreversible dada la improbabilidad estadística de que se siga el mismo proceso evolutivo dos veces, en una u otra dirección.

LEY DE HARDING-WEINBERG: En una situación evolutiva estática, un sistema de apareamiento aleatorizado conduce a una situación de equilibrio genético. La ley es completamente teórica y representa una situación estática en la que la estructura genética de la población no cambia. Describe una situación en la que no hay evolución.

LIMITACIONES: Algún tipo de fuerza que restringe o limita el movimiento de un sistema. Dicho de otra manera, restricciones impuestas por la naturaleza o por el ser humano que no permiten que sean realizadas determinadas acciones. En la práctica significa que ciertos objetivos no pueden ser alcanzados. Véase *Constraints*.

LINKAGE: La proximidad de dos o más marcadores (por ejemplo, genes) en un cromosoma; mientras más cercanos se encuentren, menor será la posibilidad de que queden separados durante los procesos de replicación o de reparación del ADN y, por lo tanto, mayor es la posibilidad de que sean heredados conjuntamente.

LOCUS: La posición de un gen u otro marcador en un cromosoma.

MAPPING: Determinación de la posición relativa de los genes en una molécula de ADN y de la distancia entre ellos.

MÁQUINAS DE BOLTZMANN: Modelos gráficos indirectos con dos variables estocásticas —en cuanto que incorporan elementos aleatorios— en los cuales los logaritmos del potencial del «clic» son funciones cuadráticas del estado de los «nodos». Se han aplicado en la computación de modelos de redes neuronales.

MÁQUINA CON UN NÚMERO FINITO DE ESTADOS: Una máquina imaginaria con un número fijo de opciones internas o posibilidades. Éstas pueden ser simplemente «sí/no» o cualquier número de posibilidades independientes, cada una de ellas determinada por una combinación de parámetros en su *input*.

MAXIMISING: Véase *Satisficing*.

MEME: Término acuñado por Richard Dawkins, por analogía con *gene* ('gen', en inglés), que define a un módulo de información contagioso que infecta y parasita la mente humana, donde se replica y altera su conducta, provocando su propagación. Los eslóganes, las frases hechas y las melodías son ejemplos de memes.

MUTACIÓN: Cambio en el genotipo que puede o no dar lugar a un cambio en el fenotipo.

NEODARWINISMO: La teoría que mantiene que la selección natural es el principal factor en la evolución y que niega específicamente la posibilidad de la herencia de caracteres adquiridos. La calificación «neo» indica que se integra la teoría de Darwin del proceso de la evolución con los modelos genéticos desarrollados a partir de Mendel y los modelos matemáticos de genética de las poblaciones.

NICHO: Un pico en la abrupta «geografía» del «paisaje adaptativo» ocupado por una criatura.

NO-LINIAL, SISTEMA: Un sistema que se comporta de modo irracional ya que no cambia proporcionalmente a los cambios en su *input*.

OPTIMISING: Véase *Satisficing*.

ORDENADOR UNIVERSAL: Un ordenador con capacidad para realizar cualquier tarea si es programado adecuadamente.

PAISAJE ADAPTATIVO: El número de nichos separados, disponibles dentro del espacio que recorre la fase de un organismo, visualizados como picos en un paisaje. Mientras más alto sea el pico mejor será la opción y mientras más suave sea la pendiente más fácil será subir hasta el pico.

PARENTESCO, COEFICIENTE DE: Se define, entre dos individuos, como el porcentaje de genes que estos individuos comparten como descendientes de un tronco común.

PLY: Número de movimientos de ajedrez que un ordenador analiza por adelantado.

POLLYANNA, SÍNDROME DE: El síndrome de una persona ciegamente optimista, denominado así por el nombre de un personaje de la literatura infantil creado por la escritora norteamericana Eleanor Porter.

POSICIÓN INTENCIONAL: Se adopta esta posición cuando un sistema es comprendido, como si fuera un agente racional, en términos de sus «creencias», sus «objetivos» y sus «intenciones».

POSICIÓN DE DISEÑO: Desde esta posición se hacen predicciones sobre un sistema de acuerdo con las características para las que fue diseñado.

POSICIÓN DE FÍSICA: Desde esta posición se consideran las clásicas leyes de la física para nuestras predicciones sobre un sistema.

PRINCIPIO ANTRÓPICO DEBIL: Las cosas que observamos están sujetas a un efecto de selección: existen en tiempos y lugares donde pueden ser percibidas por un observador.

PROCARIOTA: Organismo celular sin membrana que lo aísle del citoplasma, con un solo cromosoma. Estructura propia de bacterias y algas cianofíceas.

REDES DE HOPFIELD: Modelos neurofisiológicos utilizados para estudiar circuitos de la memoria, introducidos en la década de los ochenta por el físico John Hopfield.

REDES NEURALES: Simulación simplificada de las conexiones del cerebro humano, empleadas para la investigación del aprendizaje y la auto-organización dentro de un ambiente artificial.

RUBE GOLDBERG: Famoso personaje de cómic en Estados Unidos, inventor de máquinas imposibles.

SATISFICING: Obtener resultados bastante buenos o satisfactorios en contraste con *maximising*, que intenta buscar los resultados máximos, o con *optimising*, que busca los resultados óptimos.

SELECCIÓN DE LA PARENTELA: Se define así al mecanismo de la evolución que selecciona aquellas conductas que aumentan la adaptación inclusiva para sobrevivir.

SISTEMA: Un conjunto de partes con interrelaciones cambiantes que constituyen una totalidad integrada y consistente, aislable de su entorno.

SOMA: Todo el organismo excepto las células germinales.

STATE SPACE: El número total de posibilidades teóricas disponibles para un *sistema*, por combinación de sus partes.

TEMPLADO DE LOS METALES, SIMULACIÓN DEL: El templado de los metales es un proceso físico en el cual un sólido es calentado hasta que la temperatura alcanza un valor máximo en el cual todas las partículas del sólido se disponen aleatoriamente en la fase líquida, seguida de un lento enfriamiento. Basada en lo que sucede en sistemas naturales, la simulación del templado de los metales en los algoritmos es una estrategia de optimización para resolver problemas en organizaciones complejas.

TIPOS NATURALES: La concepción evolucionista de las especies descarta que éstas sean tipos universales y eternos, en favor de lo que el filósofo norteamericano W. V. O Quine [1969] denominó *tipos naturales*, es decir, cosas que existen naturalmente en ciertos tiempos (individuos históricos,

por lo tanto) y lugares. Las especies serían entidades biológica^ que cambian imperceptiblemente, pertenecientes, por lo tanto, a un «tipo» no reducible a una noción lógica o científicamente exigente. Esta posición de Quine ha sido denominada «naturalismo epistemológico».

TURING, MÁQUINA DE: Una forma de *ordenador universal* en la que se asume que recibe sus instrucciones de una cinta infinita de papel perforado y produce sus resultados en el mismo medio antes de parar. Como la máquina de Turing puede ser representada por una cinta, es totalmente posible suministrar una máquina de Turing como *input* a la misma máquina. Las máquinas de Turing a veces se detienen mientras que otras veces entran en un bucle infinito. Una máquina de Turing puede detenerse para una cinta de *input* y entrar en un asa infinita cuando se le introduce una cinta como *input*. El denominado «problema de la detención» plantea la siguiente cuestión: ¿es posible predecir, en general, si una máquina de Turing parará ante un *input* determinado?

UTILITARISMO: Una teoría moral, enunciada por Jeremy Bentham, según la cual una acción es correcta solamente si está conforme con el principio de la utilidad.

VIDA ARTIFICIAL: Campo científico dedicado a la comprensión de la vida mediante la abstracción de los principios dinámicos fundamentales que subyacen en los fenómenos biológicos y a la recreación de estos fenómenos dinámicos en otros medios físicos, como los ordenadores, de forma que sean accesibles a nuevos métodos de manipulación experimental.

Bibliografía

Abbot, E. A., *Flatland: A Romance in Many Dimensions* (1884), Oxford, Blackwell, 1962.

Alexander, Richard D., *The Biology of Moral Systems*, Nueva York, Gruyter, 1987.

Arbib, Michael, *Brains, Machines, and Mathematics*, Nueva York, McGraw-Hill, 1964; trad, esp.: *Cerebros, máquinas y matemáticas*, Madrid, Alianza, 1987.

—, *The Metaphorical Brain 2: Neural Networks and Beyond*, Nueva York, Wiley, 1989.

Arrhenius, S., *Worlds in the Making*, Nueva York, Harper &Row, 1908.

Ashby, Ross, *Design for a Brain*, Nueva York, Wiley, 1960; trad, esp.: *Proyecto para un cerebro*, Madrid, Tecnos, 1972.

Austin, J. L., «*A Plea for Excuses*», en Philosophical Papers, Oxford, Clarendon Press, 1961, pp. 123-152.

Axelrod, Robert, *The Evolution of Cooperation*, Nueva York, Basic Books, 1984; trad, esp.: *La evolución de la cooperación: el dilema del prisionero y la teoría de juegos*, Madrid, Alianza, 1996.

—, y William Hamilton, «The Evolution of Cooperation», *Science*, CCXI (1884), pp. 1390-1396.

Avala, Francisco J., «Beyond Darwinism? The Challenge of Macroevolution to the Synthetic Theory of Evolution», en Peter D. Asquith y Thomas Nickels, eds., *Philosophy of Science Association*, II (1982), pp. 275-291; reed, en Ruse [1989].

Ayers, M., *The Refutation of Determinism: An Essay in Philosophical Logic*, Londres, Methuen, 1968.

Babbage, Charles, *Ninth Bridgewater Treatise: A Fragment*, Londres, Murray, 1838.

Bak, Per, Henrik Flyvbjerg y Kim Sneppen, «Can We Model Darwin?», *New Scientist*, 12 de marzo de 1994, pp. 36-39.

Baldwin, J. M., «A New Factor in Evolution», *American Naturalist*, XXX (1896), pp. 441-451, 536-553.

Ball, John A., «Memes as Replicators», *Ethology and Sociobiology*, V (1984), pp. 145-161.

Barkow, Jerome H… Leda Cosmides y John Too by. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, Oxford, Oxford University Press, 1992.

Barlow, George W., y James Silverberg, eds., *Sociobiology: Beyond Nature/Nature?*, Boulder, Westview, 1980.

Baron-Cohen, Simon, *Mindblindness and the Language of the Eyes: An Essay in Evolutionary Psychology*, Cambridge, Massachusetts Institute of Technology Press, 1995.

Barret, P. H., P. J. Gautrey, S. Herbert, D. Kohn y S. Smith, eds., *Charles Darwin's Notebooks*, 1836-1844, Cambridge, British Museum, Cambridge University Press, 1987.

Barrow, J. y F. Tipler, *The Anthropic Cosmological Principle*, Oxford, Oxford University Press, 1988.

Bateson, William, «Heredity and Variation in Modern Lights», en A. C. Seward, ed., *Darwin and Modern Science*, Cambridge, Cambridge University Press, 1909, pp. 85-101.

Bedau, Mark, «Can Biological Teleology Be Naturalized?»», *Journal of Philosophy*, LXXXVIII (1991), pp. 647-657.

Bentham, Jeremy, *Introduction to the Principles of Morals and Legislation*, Oxford, Oxford University Press, 1789.

Bethell, Tom, «Darwin's Mistake», *Harpers Magazine* (febrero de 1976), pp. 70-75).

Bickerton, Derek, «The Snail Wars», *New York Times Book Review*, 3 de enero 1993, p. 5; reseña sobre Gould [1993d].

Bonner, John Tyler, *The Evolution of Culture in Animals*, Princeton, Princeton University Press, 1980; trad, esp.: *La evolución de la cultura en los animales*, Madrid, Alianza, 1982.

Borges, Jorge Luis, «La Biblioteca de Babel», en *Ficciones*, Buenos Aires, Emecé, 1944.

Brandon, Robert, «Adaptation and Evolutionary Theory», *Studies in the History and Phylosophy of Science*, IX (1978), pp. 181-206.

Breuer, Reinhard. *The Anthropic Principle: Man as the Focal Point of Nature*, Boston, Birkhäuser, 1991.

Briggs, Derek E. G., Richard A. Fortey y Matthew A. Wills, «Morphological Disparity in the Cambrian», *Science*, CCLVI (1989), pp. 1670-1673.

Brooks, Rodney, «Intelligence Without Representation», *Artificial Intelligence Journal*, XLVII (1991), pp. 139-159.

Brumbaugh, Robert M., y Rulon Wells, *The Plato Manuscripts: A New Index*, *New Haven*, Yale University Press, 1968.

Buss, Leo W., *The Evolution of Individuality*, Princeton, Princeton University Press, 1987.

Cairns-Smith, Graham, *Genetic Takeover*, Cambridge, Cambridge University Press, 1982.

—, *Seven Clues to the Origin of Life*, Cambridge, Cambridge University Press, 1985; trad, esp.: *Siete pistas sobre el origen de la vida*, Madrid, Alianza, 1990.

Calvin, William, *The River That blows Uphill: A Journey from the Big Bang to the Big Brain*, San Francisco, Sierra Club, 1986.

—, The Brain as a Darwin Machine», *Nature*, CCCXXX (1987), pp. 33-34.

Campbell, Donald, «On the Conflicts Between Biological and Social Evolution and Between Psychology and Moral Tradition», *American Psychologist* (diciembre de 1975), pp. 1103-1126.

«Comments on the Sociobiology of Ethics and Moralizing», *Behavioral Science*, XXIV (1979), PP— 37-45.

Cann, Rebecca L, Mark Stoneking y Allan C. Wilson, «Mitochondrial DNA and Human Evolution», *Nature*, CCCXXV (1987), pp. 31-36.

Capote, Truman, *In Cold Blood*, Nueva York, Random House, 1965; trad, esp.: *A sangre fría*, Barcelona, Anagrama, 1998.

Carroll, Lewis, *Through the Looking Glass*, Londres, Macmillan, 1871; trad, esp.: *A través del espejo*, Madrid, Alianza, 1998.

Changeaux, J.-P., y A. Danchin, «Selective Stabilization of Developing Synapses as a Mechanism for the Specifications of a Neuronal Networks», *Nature*, CCLXIV (1976), pp. 705-712.

Chomsky, Noam, «Three Models for the Description of Language», IRE *Transactions on Infomration Theory* IT-2 (3) (1956), pp. 13-54.

—, *Syntactic Structures*, La Hava, Mouton, 1957.

—, reseña sobre Skinner [1957], *Language*, XXXV (1959), pp. 26-58.

—, *Cartesian Linguistic*, Nueva York, Harper &Row, 1966; trad, esp.: *Lingüística cartesiana*, Madrid, Gredos, 1991.

—, *Language and Mind*, Nueva York, Harcourt Brace Jovanovich, 1972; ed. ampliada.

—, *Reflections on Language*, Nueva York, Pantheon, 1975; trad, esp.: *Reflexiones sobre el lenguaje*, Barcelona, Planeta, 1985.

—, «Rules and Representations», *Behavioral and Brain Sciences*, III (1980), pp. 1-15.
—, *Language and Problems of Knowledge: The Managua Lectures*, Cambridge, Massachusetts Institute of Technology Press, 1988.
Christensen, Scott M., y Dale R. Turner, *Folk Psychology and the Philosophy of Mind*, Hillsdale, Erlbaum, 1993.
Chruchland, Patricia S., y Terrence J. Sejnowski, *The Computational Braini Cambridge*, Massachusetts Institute of Technology Press, 1992.
Churchland, Paul, *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, Cambridge, Massachusetts Institute of Technology Press, 1989.
Clark, Andy, y Annette Karmiloff-Smith, «The Cognizer's Innards: A Psychological and Philosophical Perspective on the Development of Thought», *Mind and Language*, VIII (1994), pp. 487-519.
Clutton-Brock, T. H., y Paul H. Harvey, *Readings in Sociobiology*, San Francisco, Freeman, 1978.
Conway Morris, Simon, «Burgess Shale Faunas and the Cambrian Explosion», *Science*, CCXLVI (1989), pp. 339-346.
—, «Rerunning the Tape», *Times Literary Supplement*, 13 de diciembre de 1991, p. 6; reseña sobre Gould [1991b].
—, «Burgess Shale-type Faunas in the Context of the "Cambrian Explosion": A Review», *Journal of the Geological Society*, CXLIX (1992), pp. 631-636.
Coon, C. S., S. M. Garn y J. B. Birdsell, Races, Springfield, C. Thomas, 1950; trad, esp.: *Las razas humanas actuales*, Madrid, Guadarrama, 1969.
Cosmides, Leda, y John Tooby, «Evolutionary Psychology and the Generation of Culture», II: «Case Study: A Computational Theory of Social Exchange», *Ethology and Sociobiology*, XX (1989) pp. 51-97.
Crichton, Michael, *Jurassic Park*, Nueva York, Knopf, 1990; trad, esp: *Parque jurásico*, Barcelona, Plaza y Janés, 1997.
Crick, Francis H. C., «The Origin of the Genetic Code», *Journal of Molecular Biology*, XXXVIII (1968), p. 367.
—, *Life Itself: Its Origin and Nature*, Nueva York, Simon &C Schuster, 1981.
Crick, Francis, y Leslie E. Orgel, «Directed Panspermia», *Icarus*, XIX (1973), pp. 341-346.
Cronin, Helena, *The Ant and the Peacock*, Cambridge, Cambridge University Press, 1991.

Cummins, Robert, «Functional Analysis», *Journal of Philosophy*, LXXII (1975), pp. 741-764; reed, en Sober [1984b].

Daly, Martin, «Natural Selection Doesn't Have Goals, but It's the Reason Organisms Do», *Behavioral and Brain Sciences*, XIV (1991) pp. 219-220; comentario sobre P. J. H. Schoemaker, «The Quest for Optimality: A Positive Heuristic of Science?».

Danto, Arthur, *Nietzsche as Philosopher*, Nueva York, Macmillan, 1965.

Darwin, Charles, *On the Origin of Species by Means of Natural Selection*, Londres, Murray, 1859; trad, esp.: *El origen de las especies*, Madrid, Espasa-Calpe, 1998.

—, *Oh the Various Contrivances by Which Orchids Are Fertilised by Insects*, Londres, Murray, 1862; 2.ª ed. 1877; reimp. Chicago, Universtiy of Chicago Press, 1984.

—, *The Descent of Man, and Selection in Relation to Sex*, Londres, Murray, 1871; 2.ª ed. 1874; trad, esp.: *El origen del hombre*, Madrid, Edaf, 1987.

Darwin, Francis, *The Life and Letters of Charles Darwin*, Londres, Murray, 1887, 3 vols.; reed. Nueva York, Appleton, 1911, 2 vols.

David, Paul, «Clio and the Economics of QWERTY», *American Economic Review*, LXXV (1985), pp. 332-337.

Davies, Paul, *The Mind of God*, Nueva York, Simon &Shuster, 1992; trad, esp.: *La mente de Dios*, Barcelona, McGraw-Hill/Interamericana de España, 1993.

Dawkins, Richard, *The Selfish Gene*, Oxford, Oxford University Press, 1976; 2ª. ed. revisada, 1989; trad, esp.: *El gen egoísta: las bases biológicas de nuestra conducta*, Barcelona, Salvat, 1994.

—, «In Defence of Selfish Genes», *Philosophy*, LIV (1981), pp. 556-573.

—, *The Extended Phenotype: The Gene as the Unit of Selection*, San Francisco, Freeman, 1982.

—, «Universal Darwinism», en D. S. Bendall, ed., *Evolution from Molecules to Men*, Cambridge, Cambridge Universtiy Press, 1983, pp. 403-425.

—, «Adaptationism Was Always Predictive and Needed No Defense», *Behavioral and Brain Sciences*, VI (1983), pp. 360-361; comentario sobre Dennett (1983).

—, *The Blind Watchmaker*, Londres, Longmans, 1986; trad, esp.: *El relojero ciego*, Barcelona, Labor, 1989.

—, «Sociobiology: The New Storm in a Teacup», en Steven Rose y Lisa Appignanesi, eds., *Science and Beyond*, Oxford, Blackwell, 1986, pp. 61-78.

—, «The Evolution of Evolvability», en C. Langton, ed. *Artificial Life*, Redwood City, Addison-Wesley, 1989, vol. I, pp. 201-220.
—, reseña sobre Gould [1989a], *Sunday Telegraph*, 25 de febrero de 1990.
—, «Viruses of the Mind», en Bo Dahlbom, ed., *Dennett and His Critics*, Oxford, Blackwell, 1993, pp. 13-27.
Delius, Juan, «The Nature of Culture», en M. S. Dawkins, T. R. Hallidav y R. Dawkins, eds., *The Tinbergen Legacy*, Londres, Chapman &Hall, 1991, pp. 75-99.
Demus, Otto, *The Mosaics of San Marco in Venice*, Chicago, University of Chicago Press, 1984, 4 vols.
Dennett, Daniel C., «Machine Traces and Protocol Statements», *Behavioral Science*, XII (1968), pp. 155-161.
—, *Content and Consciousness*, Londres, Rout ledge 8c Kegan Paul, 1969; trad, esp.: *Contenido y conciencia*, Barcelona, Gedisa, 1996.
—, «The Abilities of Men and Machines» (1970), en *Brainstorms*, Cambridge, Massachusetts Institute of Technology Press/A Bradford Book, 1978, pp. 256-266.
—, «Intentional Systems», *Journal of Philosophy*, LXVIII (1971), pp. 87-106.
—, reseña sobre Lucas [1970], *Journal of Philosophy*, LXIX (1972), pp. 527-531.
—, «Why the Law of Effect Will Not Go Asay», *Journal of the Theory of Social Behaviour*, V (1975), pp. 179-187; reed, en *Brainstorms*, Cambridge, Massachusetts Institute of Technology Press/A Bradford Book, 1978.
—, *Brainstorms*, Cambridge, Massachusetts Institute of Technology Press/A Bradford Book, 1978.
—, «Passing the Buck to Biology», *Behavioral and Brain Sciences*, III (1980), p. 19.
—, «Three Kinds of Intentional Psychology», en R. Healey, ed., *Reduction, Time and Reality*, Cambridge, Cambridge University Press, 1981, pp. 37-61.
—, «Intentional Systems in Cognitive Ethology: The “Panglossian Paradigm” Defended», *Behavioral and Brain Sciences*, VI (1983), pp. 343-390.
—, *Elbow Room: The Varieties of Free Will Worth Wanting*, Cambridge, Massachusetts Institute of Technology Press, 1984.
—, «Can Machines Think?», en M. Shafto, ed., *How We Know*, San Francisco, Harper &Row, 1985, pp. 121-145.

—, «The Logical Geography of Computational Approaches: A View from the East Pole», en M. Brand y M. Harnish, eds., *Problems in the Representation of Knowledge*,Tucson, University of Arizona Press, 1987, pp. 59-79.
—, *The Intenttional Stance*, Cambridge, Massachusetts Institute of Technology Press/A Bradford Book, 1987.
—, «When Philosophers Encounter Artificial Intelligence», *Daedalus*, CXVII (1988), pp. 283-295.
«The Moral First Aid Manual»», en Sterling M. McMurrin, ed., *Tanner Lectures on Human Values*, Salt Lake City, University of Utah Press, 1988, vol. VIII, pp. 120-147.
—, reseña sobre Robert J. Richards [1987], *Philosophy of Science*, LVI, 3 (1989), pp. 540-543.
—, «Murmurs in the Cathedral», *Times Literary Supplement*, 26 de septiembre-5 de octubre de 1989, pp. 1066-1068; reseña sobre Penrose (1989).
—, «Teaching an Old Dog New Tricks», *Behavioral and Brain Sciences*, XIII (1990), pp. 76-77; comentario sobre Schull (1990).
—, «The Interpretation of Texts, People, and Other Artifacts», *Philosophy and Phenomenological Research*, L (1990), pp. 177-194.
—, «Memes and the Exploitation of Imagination», *Journal of Aesthetics and Art Criticism*, XLVIII (1990), pp. 137-135.
—, *Consciousness Explained*, Boston, Little Brown, 1991; trad, esp.: *La conciencia explicada: una teoría interdisciplinar*, Barcelona, Paidós, 1995.
—, «Real Patterns», *Journal of Philosophy*, LXXXVII (1991), pp. 27-51.
—, «Granny’s Campaign for Safe Science», en B. Loewer y G. Rey, eds., *Meaning in Mind: Eodor and His Critics*, Oxford, Blackwell, 1991, pp. 87-94.
—, «The Braind and Its Boundaries», *Times Literary Supplement*, 10 de mayo de 1991; reseña sobre McGinn [1991].
—, «Ways of Establishing Harmony», en B. McLaughlin, ed., *Dretske and His Critics*, Oxford, Blackwell, 199T; reed, revisada en E. Villanueva, ed. *Information, Semantics, and Epistemology*, México, Sociedad Filosófica Iberoamericana/Oxford, Blackwell, 1991.
—, «La comprehension artisanale», en Denis Eisette, ed., Daniel C. *Dennett et les stratégies intentionnelles*, *Lekton*, XI (invierno de 1992), pp. 27-52; trad, francesa de «Do-It-Yourself Understanding».

—, «Down with School! Up with Logoland!», *New Scientist*, 6 de noviembre de 1993, pp. 45-46; reseña sobre Papcrt [1993].
—, «Confusion over Evolution: An Exchange», *New York Review of Books*, 14 de enero de 1993, pp. 43-44.
—, reseña sobre John Searle (1992), *Journal of Philosophy*, XC O993), pp. 193-205.
—, «Living on the Edge», *Inquiry*, XXXVI (1993), pp. 135-159.
—, «Cognitive Science as Reverse Engineering: Several Meanings of "Top-down" y "Bottom-up"», en D. Prawitz, B. Skyrms y D. Westerstahl, eds., *Proceedings of the youth International Congress of Logic, Methodology and Philosophy of Science*, Amsterdam, North-Holland, 1994.
—, «Language and Intelligence», en Jean Khalfa, ed., *What Is Intelligence*?, Cambridge, Cambridge University Press, 1994, pp-161-178.
—, «Labeling and Learning», *Mind and Language*, VIII (1994)» PP-540-548; comentario sobre Clark y Karmiloff-Smith [1994].
—, «E Pluribus Unum?», *Behavioral and Brain Sciences*, XVII (1994) pp. 617-618.
—, «The Practical Requirements for Making a Conscious Robot», *Proceedings of the Royal Society* (1994).
—, y John Haugeland, «Intentionality», en R. L. Gregory, ed., *The Oxford Companion to the Mind*, Oxford, Oxford University Press, 1987, pp. 383-386.
Denton, Michael, *Evolution: A Theory in Crisis*, Londres, Burnett, 1985.
Descartes, René, *Discurso del método (Discours de la méthode*, Leyden, 1637).
Desmond, Adrian, y James Moore, *Darwin*, Londres, Michael Joseph, 1991.
Deutsch, D., «Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer», *Proceedings of the Royal Society*, CD (1985), pp. 97-117.
De Vries, Peter, *The Vale of Laughter*, Boston, Little Brown, 1953.
Dewdney, A. K., T*he Planiverse*, Nueva York, Poseidon, 1984.
Dewey, John, *The Influence of Darwin on Philosophy*, Nueva York, Holt, 1910; reed. Bloomington, Indiana Unviersity Press, 1965.
Diamond, Jared, «The Saltshaker's Curse», *Natural History* (octubre de 1991), pp. 20-26.
—, *The Third Chimpanzee: The Evolution and Future of the Human Animal*, Nueva York, HarperCollins, 1992; trad, esp.: *El tercer chimpancé*, Madrid, Espasa-Calpe, 1994.

Diderot, Denis, *Lettre sur les aveugles, à l'usage de ceux qui voyent*, Londres, 1749; trad, esp.: *Carta sobre los ciegos para uso de los que ven*, Madrid, Piqueta, 1978

Dietrich, Michael, «Macromutation», en Evelyn Fox Keller y Elisabeth A. Lloyd, eds., *Keywords in Evolutionary Biology*, Cambridge, Harvard University Press, 1992, pp. 194-201.

Diggins, John Patrick, reseña sobre Marsden, *The Soul of the American University*, New York Times Book Review, 17 de abril de 1994, P. 25.

Dobzhansky, Theodosius, «Nothing in Biology Makes Sense Except in the Light of Evolution», *American Biology Teacher*, XXXV (1973) pp. 125-129.

Donald, Merlin, *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*, Cambridge, Harvard University Press, 1991.

Doolittle, W. F., y C. Sapienza, «Selfish Genes, the Phenotype Paradigm and Genome Evolution», *Nature*, CCLXXXIV (1980), pp. 601-603.

Dretske, Fred, «Misrepresentation», en R, Bogdan, ed., *Belief*, Oxford, Oxford University Press, 1986.

Dreyfus, Hubert, «Alchemy and Artificial Intelligence», RAND Technical Report P-3244 (diciembre de 1965).

—, *What Computers Can't Do: The Limits of Artificial Intelligence*, Nueva York, Harper &Row, 1972; reed, revisada, 1979.

Dyson, Freeman, *Disturbing the Universe*, Nueva York, Harper &Row, 1979.

Eccles, John, *The Neurophysiological Basis of Mind*, Oxford, Clarendon, 1953.

Eckert, Scon A., «Bound for Deep Water», *Natural History* (marzo de 1992), pp. 28-35.

Edelman, Gerald, *Neural Darwinism*, Nueva York, Basic Books, 1987.

—, *Bright Air, Brilliant Fire*, Nueva York, Basic Books, 1992.

Edwards, Paul, «Professor Tillich's Confusions», *Mind*, LXXIV (1965), pp. 191-214.

Eigen, Manfred, «Wie entsteht Information? Prinzipien der Selbstorganisation in der Biologie», *Berichteet der Bunsengesellschaft für Physikalische Chemie*, LXXX (1976), p. 1059.

—, «Self-Replication and Molecular Evolution», en D. S. Bendall, ed., *Evolution from Molecules to Men*, Cambridge, Cambridge University Press, 1983, pp. 105-130.

—, *Steps towards Life*, Oxford, Oxford University Press, 1992.

—, y R. Winkler-Oswatitsch, Das Piel, Munich, 1975; trad, inglesa: *Laws of the Game*, Nueva York, Knopf, 1981.

—, y P. Schuster, 1977. «The Hypercycle: A Principle of Natural Self-Organization. Part A: Emergence of the Hypercycle», *Naturwissenschaften*, LXIV (1975), pp. 541-565.

Eldredge, Niles, «A la recherche du Docteur Pangloss», *Behavioral and Brain Sciences*, VI (1983), pp. 361-362; comentario sobre Dennett (1983).

—, *Times Frames: The Rethinking of Darwinian Evolution and the Theory of Punctuated Equilibria*, Nueva York, Simon &Schuster, 1985.

—, *Macroevolutionary Dynamics: Species, Niches and Adaptive Peaks*, Nueva York, McGraw-Hill, 1989.

—, y S. J. Gould, «Punctuated Equilibria: An Alternative to Phyletic Gradualism», en T. J. M. Schopf, ed., *Models in Paleobiology*, San Francisco, Freeman, Cooper and Company, 1972, pp. 82-115; reed, en Edredge (1985:193-223).

Ellegard, Alvar, «The Darwinian Theory and the Argument from Design», *Lychnos* (1956), pp. 173-192.

—, *Darwin and the General Reader*, Goteborg, Goteborg University Press, 1958.

Ellestrand, Norman, «Why Are Juveniles Smaller Than Their Parents?», *Evolution*, XIII (1983), pp. 1091-1094.

Ellis R. J., y S. M. van der Vies, «Molecular Chaperones», *Annual Review of Biochemistry*, LX (1991), pp. 321-347.

Elsasser, Walter, *The Physical Foundations of Biology*, Oxford, Oxford University Press, 1958.

—, *Atom and Organism*, Princeton, Princeton University Press, 1966.

Engels, W. R., «The Origin of P Elements in “Drosophila melanogaster”», *BioEssays*, XIV (1992), pp. 681-686.

Ereshefsky, Marc, ed., *The Units of Evolution: Essays on the Nature of Species*, Cambridge, Massachusetts Institute of Technology Press/A Bradford Book, 1992.

Eshel, I., «Are Intragenetic Conflicts Common in Nature? Do They Represent an Important Factor in Evolution?», *Journal of Theoretical Biology*, CVIII (1984), pp. 159-162.

—, «Evolutionary Generic Stability of Mendelian Segregation and the Role of Free Recombinatin in the Chromosomal System», *American Naturalist*, CXXV (1985), pp. 412-420.

Fausto-Sterling, Anne, *Myths of Gender: Biological Theories About Women and Men*, Nueva York, Basic Books, 1985; 2.ª ed. 1992.

Feduccia, Alan, «Evidence from Claw Geometry Indicating Arboreal Habits of “Archeopteryx”», *Science*, CCLIX (1993), pp. 790-793.

Feigenbaum, F. A., y J. Feldman, *Computers and Thought*, Nueva York, McGraw-Hill, 1964.

Feynman, Richard, *What do You Care What Other People Think*?, Nueva York, Bantam, 1988; trad, esp.: *¿Qué te importa qué piensen los demás?*, Madrid, Alianza, 1990.

Fisher, Dan, «Swimming and Burrowing in “Limulus” and “Mesolimulus”», *Fossils and Strata*, IV (1975), pp. 281-290.

Fisher, R. A., *The Genetical Theory of Natural Selection*, Oxford, Clarendon, 1950.

Kitchen, John, *The Construction fo Gothic Cathedrals*, Oxford, Clarendon, 1961.

—, *Building Construction Before Mechanization*, Cambridge, Massachusetts Institute of Technology Press, 1986.

Fodor, Jerry, *The Language of Thought*, Hassocks, Harvester, 1975; trad, esp.: *El lenguaje del pensamiento*, Madrid, Alianza, 1985.

—, «Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology», *Behavioral and Brain Sciences*, III (1980), pp. 63-110.

—, *The Modularity of Mind*, Cambridge, Massachusetts Institute of Technology Press, 1983; trad, esp.: *La modularidad de la mente*, Madrid, Morata, 1994.

—, *Psychosemantics*, Cambridge, Massachusetts Institute of Technology Press, 1987; trad, esp.: *Psicosemántica*, Madrid, Tecnos, 1994.

—, *A Theory of Content and Other Essays*, Cambridge, Massachusetts Institute of Technology Press, 1990.

—, «The Big Idea: Cant There Be a Science of Mind?», *Times Literary Supplement*, 3 de julio de 1992, p. 5.

Foote, Mike, «Cam hi ran and Recent Morphological Dispa rity». *Science*, CCLVI (1992.). p. 1670; respuesta a Briggs y otros (1989).

Forbes, Graeme, «Thisness and Vagueness», *Synthese*, LIV (1983), pp. 235-259.

—, «Two Solutions to Chisholm’s Paradox», *Philosophical Studies*, XLVI (1984) pp. 171-187.

Fox, S. W., y K. Dose, *Molecular Evolution and the Origin of Life*, San Francisco, Freeman, 1972.
Futuyma, Douglas, *Science on Trial: The Case for Evolution*, Nueva York, Pantheon, 1982.
Gabbey, Allan, «Descartes, Newton and Mechanics: The Disciplinan' Turn», Tufts Philosophy Colloquium, noviembre de 1993.
Galilei, Galileo, *Dialogo sopra i due massimi sistemi del mondo tolemaico e copernicano*, Florencia, 1632; trad, esp.: *Diálogo sobre los dos máximos sistemas del mundo ptolemaico y copernicano*, Madrid, Alianza, 1995.
Gardner, Martin, «Mathematical Games», *Scientific American*, CCXXIII (octubre de 1970), pp. 120-123.
—, «Mathematical Games», *Scientific American*, CCXXIV (febrero de 1971), pp. 112-17.
—, «WAP, SAP, PAP and FAP», *New York Review of Books*, mayo de 1986; reed, ampliada en Martin Gardner, *Gardner's Whys and Wherefores*, Chicago, University of Chicago Press, 1989.
Gauthier, David, *Morals by Agreement*, Nueva York, Oxford University Press, 1986; trad, esp.: *La moral por acuerdo*, Barcelona, Gedisa, 1994.
Gee, Henry, «Something Completely Different», *Nature*, CCCLVIII (1992), pp. 456-457.
Gert, Bernard, *The Moral Rules*, Nueva York, Harper &Row, 1966; reed. Nueva York, Harper Torchbook, 1973.
Ghiselin, M., «Lloyd Morgan's Canon in Evolutionary Context», *Behavioral and Brain Sciences*, VI (1983), pp. 362-363.
Gibbard, Alan, «Moral Judgment and the Acceptance of Norms», y «Reply to Sturgeon», *Ethics*, XCVI (1985), pp. 5-41.
Gilkey, Langdon, *Creationism on Trial: Evolution and God at Little Rock*, San Francisco, Harper &Row, 1985.
Gille, Bertrand, *Engineers of the Renaissance*, Cambridge, Massachusetts Institute of Technology Press, 1966.
Gingerich, Philip, «Rate of Evolution: effects of Time and Temporal Scaling», *Science*, CCXXII (1983), pp. 159-161.
—, Respuesta a Gould [1983c], *Science*, CCXXVI (1984), pp. 995-996.
Gjertsen, Derek, *Science and Philosophy: Past and Present*, Londres, Penguin, 1989.
Gödel, Kurt, «Über formal Unentscheidbare Sátze der *Principia Mathematica* und verwande Systeme, I», *Monatshefte für Mathematik und Physik*, XXXVIII (1931), pp. 173-198.

Godfrey-Smith, Peter, «Spencerian Explanation and Constructivism», 5 de noviembre de 1993; coloquio de filosofía del Massachusetts Institute of Technology.
Goldschmidt, Richard B., «Some Aspects of Evolution», *Science*, LXXVII1 (1933), pp. 539-547.
—, *The Material Basis of Evolution*, Seattle, University of Washington Press, 1940.
Goodman, Nelson. *Fact, Fiction and Forecast*, 2.ª ed. Nueva York, Bobbs-Merrill, 1965.
Goodwin, Brian, «Is Biology an Historical Science?», en Steven Rose y Lisa Appignanesi, eds., *Science and Beyond*, Oxford, Blackwell, 1986, pp. 47-60.
Gould, Stephen Jay, *Ever Since Darwin*, Nueva York, Norton, 1977; trad, esp.: *Desde Darwin (Reflexiones sobre historia natural)*, Barcelona, Blume, 1983.
—, *Ontogeny and Phytogeny*, Cambridge, Belknap, 1977.
—, *The Panda's Thumb*, Nueva York, Norton, 1980; trad, esp.: *El pulgar del panda*, Barcelona, Crítica, 1994.
—, *«Is a New and General Theory of Evolution Emerging?», Paleology*, VI (1980), pp. 119-130.
—, «Sociobiology and the Theory of Natural Selection», *American Association for the Advancement of Science Symposia*, XXXV (1980), pp. 157-169; reed, en Ruse (1989).
—, «The Evolutionary Biology of Constraint», *Daedalus*, CIX (1980), pp. 39-52.
—, *The Mismeasure of Man*, Nueva York, Norton, 1981; trad, esp.: *La falsa medida del hombre*, Barcelona, Crítica, 1997.
—, «Darwinism and the Expansion of Evolutionary Theory», *Science*, CCXVI (1981), pp. 380-387; reed, en Ruse (1989).
—, «The Uses of Heresy: An Introduction to Richard Goldschmidt's *The Material Basis of Evolution*», en R. Goldschmidt, *The Material Basis of Evolution*, Hew Haven, Yale University Press, 1982.
—, «The Meaning of Punctuated Equilibrium, and Its Role in Validating a Hierarchical Approach to Macroevolution», en R. Milkman, ed., *Perspectives on Evolution*, Sunderland, Sinauer, 1981, pp. 83-104.
«Change in Developmental Timing as a Mechanism of Macroevolution», en J. T. Bonner, ed., *Evolution and Development*, Berlin, Heidelberg/Nueva York, Springer-Verlag, 1981.

«The Hardening of the Modern Synthesis», en M. Grene, ed., *Dimensions of Darwinism*, Cambridge, Cambridge University Press, 1983, pp. 71-93.

Hens Teeth and Horse's Toes, Nueva York, Norton, 1983; trad, esp.: *Dientes de gallina, dedos de caballo*, Barcelona, Blume, 1984.

—, «Smooth Curve of Evolutionary Rate: A Psychological and Mathematical Artifact», *Science*, CCXXVI (1983), pp. 994-995.

—, *The Flamingo's Smile*, Nueva York, Norton, 1985; trad, esp.: *La sonrisa del flamenco: reflexiones sobre historia natural*, Barcelona, Crítica, 1995.

«Darwinism Defined: The Difference Between Fact and Theory», *Discover* (enero de 1987)» pp. 64-70.

—, *Wonderful Life: The Burgess Shale and the Nature of History*, Nueva York, Norton, 1989; trad, esp.: *La vida maravillosa*, Barcelona, Crítica, 1995.

—, «Tires to Sandals», *Natural History* (abril de 1989), pp. 8-15.

—, *The Individual in Darwin's World*, Edinburgh, Edinburgh University Press, 1990.

—, «The Panda's Tumb of Technology» (1991) en Gould [1991b:59-75].

—, *Bully for Brontosaurus*, Nueva York, Norton, 1991; trad, esp.: *Brontosaurus y la nalga del ministro*, Barcelona, Crítica, 1993.

—, «The Confusion over the Evolution», *New York Review of Books*, 19 de noviembre de 1992, pp. 47-54.

—, «Life in a Punctuation», *Natural History*, CI (octubre de 1992), pp. 10-21.

—, «Fulfilling the Spandrels of World and Mind», en Selzer [1993:310-336].

—, ed., *The Book of Life*, Nueva York, Norton, 1993; trad, esp.: *El libro de la vida*, Barcelona, Crítica, 1993.

—, «Cordelia's Dilemma», *Natural History*, CIII (febrero de 1993), pp. 10-19.

—, *Eight Little Piggies*, Nueva York, Norton, 1993; trad, esp.: *Ocho cerditos*, Barcelona, Crítica, 1994.

—, «Confusion over Evolution: An Exchange», *New York Review of Books*, 14 de enero de 1993, pp. 43-44.

—, y N. Eldredge, «Puntuated Equilibrium Comes of Age», *Nature*, CCCLXVI (1993), pp. 223-227.

—, y R. Lewontin, «The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme», *Proceedings of the Royal Society*, CCV (1979), pp. 581-598.

—, y Elizabeth Vrba, «Exaptation: A Missing Term in the Science of Form», *Paleobiology*, VIII (1981), pp. 4-15.

Greenwood, John D., ed., *The Future of Folk Psychology: Intentionality and Cognitive Science*, Cambridge, Cambridge University Press, 1991.

Gregory, R. L., *Mind in Science: A History of Explanations in Psychology and Physics*, Cambridge, Cambridge University Press, 1981.

—, ed., *The Oxford Companion to the Mindy Oxford*, Oxford University Press, 1987.

Grice, H. P., «Meaning», *Philosophical Review*, LXV1 (1957), pp. 377-388.

—, «Utterers Meaning and Intentions», *Philosophical Review*, LXXVIII (1969), pp. 147-177.

Grossberg, Stephen, «Adaptive Pattern Classification and Universal Recording: Part I. Parallel Development and Coding of Neural Feature Detectors», *Biological Cybernetics*, XXIII (1976)» pp. 121-134.

Hadamard, Jacques, *The Psychology of Inventing in the Mathematical Field*, Princeton, Princeton University Press, 1949.

Haig, David, «Genomic Imprinting and the Theory of Parent-Offspring Conflict», *Developmental Biolog*y, III (1992), pp. 153-160.

—, «Genetic Conflicts in Human Pregnancy», *Quarterly Review of Btology*, LXVIII (1993), pp. 495-532.

—, y A. Grafen, «Genetic Scrambling as a Defence Agaist Meiotic Drive», *Journal of Theoretical Biology*, CLIII (1991), pp. 531-558.

—, y Chris Graham, «Genomic Imprinting and the Strange Case of the Insulin-like Growth Factor II Receptor», *Cell*, LXIV (1991), pp. 1045-1046.

—, y M. Westoby, «Parent-specific Gene Expression and the Triploid Endosperm», *American Naturalist*, CXXXIV (1989), pp. 147-155.

Hale, Ken, y otros, «Endangered Languages», *Language*, LXVIII (1992), pp. 1-42.

Hamilton, William, «The Art of Publishing Obscurely», en G. Hardin, ed., *Population, Evolution and Birth Control: A Collection of Controversial Reading*, San Francisco, Freeman, 1964, pp. 116-119.

—, «The Tragedy of the Commons», *Science*, CLXII (1968), pp. 1243-1248.

Hardy, A lister, «Was Man More Aquatic in the Past?», *New Scientist*, 1960, pp. 642-645.

Haugeland, John, *Artificial intelligence: The Very Idea*, Cambridge, Massachusetts Institute of Technology Press, 1985.

Hawking, Stephen W., *A Brief History of Time*, Nueva York, Bantam, 1988; trad, esp.: *Historia del tiempo*, Barcelona, Grijalbo, 1998.

Hebb, Donald, *The Organization of Behavior*, Nueva York, Wiley, 1949.

Hinton, Geoffrey E., y S. J. Nowlan, «How Learning Can Guide Evolution», en *Complex Systems*, Carnegie-Mellon University, 1987, vol. I, pp. 495-502.

Hobbes, Thomas, *Leviatán* (*Leviathan*, Londres, 1651).

Hodges, Andrew, *Alan Turing: The Enigma*, Nueva York, Simon &Schuster, 1983.

Hofstadter, Douglas, *Gödel Escher Bach*, Nueva York, Basic Books, 1979; trad, esp.: *Gödel, Escher, Bach*, Barcelona, Tusquets, 1989.

—, *Metamagical Themas: Questing for the Essence of Mind and Pattern*, Nueva York, Basic Books, 1985.

—, y Daniel C. Dennett, *The Mind's I*, Nueva York, Basic Books, 1981.

Holland, John, *Adaptation in Natural and Artificial Systems*, Ann Arbor, University of Michigan Press, 1975.

—, «Complex Adaptive Systems», *Daedalus* (invierno de 1992), p. 25.

Hollingdale, R. J., *Nietzsche: The Man and His Philosophy*, Londres, Routledge &Kegan Paul, 1965.

Houck, Marilyn A., Jonathan B. Clark, Kenneth R. Peterson y Margaret G. Kidwell, «Possible Horizontal Transfer of "Drosophila" Genes by the Mite "Procrolaelaps Regalis"», *Science*, CCLIII [1991], pp. 1125-1129.

Houston, Alasdair, «Matching, Maximizing, and Melioration as Alternative Descriptions of Behaviour», en J. A. Meyer y S. Wilson, eds., *From Animals to Animals*, Cambridge, Massachusetts Institute of Technology Press, 1990, pp. 498-509.

Hoy, David, «Nietzsche, Hume, and the Genealogical Method», en Y. Yovel, ed., *Nietzsche as Affirmative Thinker*, Dordrecht, Martinus Nijhoff, 1986, pp. 20-38.

Hoyle, Fred, *Of Men and Galaxies*, Seattle, University of Washington Press, 1964.

—, y Chandra Wickramasinghe, *Evolution from Space*, Londres, Dent, 1981.

Hrdy, Sarah Blaffer, *The Langurs of Aim: Female and Male Strategies of Reproduction*, Cambridge, Harvard University Press, 1977.

Hull, David, «Individuality and Selection», *Annual Review of Ecology and Systematics*, XI (1980), pp. 311-332.

—, «The Naked Meme», en H. C. Plotkin, ed., *Learning, Development and Culture*, Nueva York, Wiley, 1982, pp. 273-327.

Hume, David, *Tratado de la naturaleza humana* (*A Treatise of Human Nature*, 1739).

—, *Dialogues Concerning Natural Religion*, Londres, 1779; trad, esp.: *Diálogos sobre la religión natural*, Madrid, Tecnos, 1994.

Humphrey, Nicholas, «The Social Function of Intellect», en P. P. G. Bateson y R. A. Hinde, eds., *Growing Points in Ethology*, Cambridge, Cambridge University Press, 1976, pp. 303-317.

—, «The Adaptiveness of Mentalism?», *Behavioral and Brain Sciences*, III (1983), p. 366.

—, *The Inner Eye*, Londres, Faber &Faber, 1986; trad, esp.: *La mirada interior*, Madrid. Alianza, 1995.

—, «Scientific Shakespeare», *The Guardian*, 26 de agosto de 1987.

Isack, H. A., y H. U. Reyer, «Honeyguides and Honey Gatherers: Interspecific Communication in a Symbiotic Relationship», *Science*, CCXLIII (1989), pp. 1343-1346.

Israel, David, *The Role of Proposiotional Objects of Belief in Action*, CSLI Monograph Report CSLI-87-72. Palo Alto, Stanford University Press, 1987.

Jackendoff, Ray, *Consciousness and the Computational Mindy Cambridge*, Massachusetts Institute of Technology Press/A Bradford Book, 1987.

—, *Patterns in the Mind: Language and Human Nature*, Londres, Harvester Wheatsheaf, 1993.

Jacob, François, *The Possible and the Actual*, Seattle, University of Washington Press, 1982.

—, «Molecular Tinkering in Evolution», en D. S. Bendall, ed., *Evolution from Molecules to Men*. Cambridge, Cambridge University Press, 1983, pp. 131-144.

James, William, *Lecture Notes* 2880-1897, 1880; cit. en Sills y Merton [1991].

Jaynes, Julian, *The Origins of Consciousness in the Breakdown of the Bicameral Mind*, Boston, Houghton Mifflin, 1976.

Jones, Steve, «A Slower Kind of Bang», *London Review of Books*, abril de 1993, p. 20; reseña sobre E. O. Wilson, *The Diversity of Life*.

Kant, Immanuel, *Fundamentación de la metafísica de las costumbres* (Grundlegung zur Metaphysik der Sitien, Königsberg, 1785).

Kauffman, Stuart, *The Origins of Order: Self-Organization and Selection in Evolution*, Nueva York, Oxford University Press, 1993.

Kaufmann, Walter, *Nietzsche: Philospher, Psychologist, Antichrist*, Princeton, Princeton University Press, 1950; reimp. Nueva York, Meridian Paperback, 1956.

Keil, Frank, C., «The Origins of an Autonomous Biology», en M. Gunnar y M. Maratsos, eds., *Modularity and Constraints in Language and Cognition: The Minnesota Symposia on Child Psychology*, Hillsdale, Erlbaum, 1991, vol. XXV, pp. 103-137.

Keller, Evelyn Fox, y Elisabeth A. Lloyd, eds., *Keywords in Evolutionary Biology*, Cambridge, Harvard University Press, 1992.

King's College Sociobiology Group, *Current Problems in Sociobiology*, Cambridge, Cambridge University Press, 1982.

Kipling, Rudyard, *Just So Stories*, 1911; reimp. Garden City, Doubleday, 1952; trad, esp.: *Los cuentos de así fue…*, Madrid, Akal, 1987.

Kirkpatrick, S., C. D. Gelatt y M. P. Vecchi, «Optimization by Simulated Annealing», *Science*, CCXX (1983), pp. 671-680.

Kitcher, Philip, *Abusing Science*, Cambridge, Massachusetts Institute of Technology Press, 1982.

—, «Species», *Philosophy of Science*, LI (1984), pp. 308-333.

—, «Darwin's Achievement», en N. Rescher, ed., *Reason and Rationality in Science*, Lanham, University Press of America, 1985, pp. 127-189.

Vaulting Ambition, Cambridge, Massachusetts Institute of Technology Press, 1985.

«The Evolution of Human Altruism», *Journal of Philosophy*, XC (1993) pp. 497-516.

Krautheimer, *Richard, Early Christian and Byzantine Architecture*, 3.ª ed. Londres, Penguin, 1981; trad, esp.: *Arquitectura paleocristiana y bizantina*, Madrid, Cátedra, 1988.

Krebs, John R., y Richard Dawkins, «Animal Signals: Mind-Reading and Manipulation», en J. R. Krebs y N. B. Davies, eds., *Behavioural Ecology: An Evolutionary Approach*, 2.ª ed. Oxford, Blackwell, 1984, pp, 380-402.

Küppers, Bernd-Olaf, *Information and the Origin of Life*, Cambridge, Massachusetts Institute of Technology Press, 1990.

Lancaster, Jane, *Primate Behavior and the Emergence of Human Culture*, Nueva York, Holt, Rinehart and Winston, 1975.

Landman, Otto E., «The Inheritance of Acquired Characteristics», *Annual Review of Genetics*, XXV (1991), pp. 1-20.

—, «Inheritance of Acquired Characteristics», *Scientific American* (marzo de 1993), p. 150.

Langton, Christopher, Charles Taylor, J. Doyne Farmer y Steen Rasmussen, *Artificial Life II*, Redwood City, Addison-Wesley, 1992.

Leibniz, Gottfried Wilhelm, *Teodicea (Essais de Théodicée sur la bonté de Dieu, la liberté de l'homme et l'origine du mal*, Amsterdam, 1710).
Leigh, E. G., *Adaptation and Diversity*, San Francisco, Freeman, Cooper and Company, 1971».
Lenat, Douglas B., y R. V. Guha, Building Large Knowledge-based Systems: Representation and Inference in the CYC Project, Reading, Addison-Wesley, T990.
Leslie, Alan, «Pretense, Autism and the Theory-of-Mind Module», *Current Directions in Psychological Science*, I (T992), pp. 18-21.
Leslie, John, Universes, Londres, Routledge &Kegan Paul, 1989.
Lettvin, J. Y., U. Maturana, W. McCulloch y W. Pitts, «What the Frogs Eye Tells the Frog's Brain», en *Proceedings of the IRE*, 1959» pp. 1940-1951.
Levi, Primo, *Il sistema periodico*, Roma, Teknos, 1984; trad, esp.: *El sistema periódico*, Madrid, Alianza, 1998.
Lévi-Strauss, Claude, *La pensée sauvage*, Paris, Plon, 1962.
Lewin, Roger, *Complexity: Life at the Edge of Chaos*, Nueva York, Macmillan, 1992; trad, esp.: *Complejidad: el caos como orden generador del orden*, Barcelona, Tusquets, 1995.
Lewis, David, *Philosophical Papers*, Oxford, Oxford University Press, 1986, vol. II.
Lewontin, Richard, «Adaptation», en *The Encyclopedia Einaudi*, Milán, Einaudi, 1980.
—, «Elementary Errors About Evolution», *Behavioral and Brain Sciences*, VI (1983), pp. 367-368; comentario sobre Dennett (1983).
—, «The Shape of Optimality», en John Dupré, ed., *The Latest on the Best: Essays on Evolution and Optimality*, Cambridge, Massachusetts Institute of Technology Press, 1987.
—, Steven Rose y Leon Kamin, *Not in our Genes: Biology Ideology and Human Nature*, Nueva York, Pantheon, 1984; trad, esp.: *No está en los genes*, Barcelona, Crítica, 1987.
Linnaeus, Carolus, *Philosophia Botánica*, 1751.
Lloyd, M., y H. S. Dybass, «The Periodical Cicada Problem», *Evolution*, XX (1966), pp. 132-149.
Locke, John, *Ensayo sobre el entendimiento humano (Essay Concerning Human Understandings*, Londres, 1690).
Lord, Albert, *The Singer of Tales*, Cambridge, Harvard University Press, 1960.

Lorenz, Konrad, *Die Rückseite des Spiegels*, Munich, R. Piper-Verlag, 1973; trad, esp.: *Otra cara del espejo*, Barcelona, Plaza y Janés, 1985.
Lovejoy, Arthur O., *The Great Chain of Being: A Study of the History of an Idea*, Nueva York, Harper &Row, 1936; trad, esp.: *La gran cadena del ser*, Barcelona, Icaria, 1983.
Lucas, J. R., «Minds, Machines, and Gödel», *Philosophy*, XXXVI (1961), pp. 1-12.
—, *The Freedom of the Will*, Oxford, Oxford University Press, 1970.
MacKenzie, Robert Beverley, *The Darwinian Theory of the Transmutation of Species Examined*, Cambridge, Nisbet and Co., 1868; cit. en Athenaeum, MMCII (febrero de 1868), p. 217.
Malthus, Thomas, *Essay on the Principle of Population*, Londres, J. Johnson, 1798; trad, esp.: *Ensayo sobre el principio de la población*, Madrid, Akal, 1990.
Margolis, Howard, Patterns, *Thinking and Cognition*, Chicago, University of Chicago Press, 1987.
Margulis, Lynn, *Symbiosis in Cell Evolution*, San Francisco, Freeman, 1981.
—, y Dorion Sagan, 1986, *Microcosmos*, Nueva York, Simon 6c Schuster, 1981.
—, «Bacterial Bedfellows», *Natural History*, XCVI (marzo de 1987), pp. 16-33.
Marks, Jonathan, reseña sobre Diamond [1992], *Journal of Human Evolution*, XXIV (1993), pp. 69-73.
—, «Scientific Misconduct: Where “Just Say No” Fails», *American Scientist*, LXXXI (julio-agosto de 1993), pp. 380-381.
Martin, J., M. May hew, T. Langer y F. U, Hard, «The Reaction Cycle of GroEL and GroES in Chaperonin-assisted Protein Folding», *Nature*, CCCLXVI (1993), pp. 118-133.
Masserman, Jules H., Stanley Wechkin y William Terris, «“Altruistic” Behavior in Rhesus Monkeys», *American Journal of Psychiatry*, CXXI (1964), pp. 584-585.
Matthew, Patrick, *Naval Timber and Arboriculture*, 1831.
—, Carta al director. *Gardener Chronicle*, 7 de abril de 1860.
Maynard Smith, John, *The Theory of Evolution*, Cambridge, Cambridge University Press, 1958.
—, *On Evolution*, Edinburgh, Edinburgh University Press, 1971; trad, esp.: *Acerca de la evolución*, Barcelona, Blume, 1979.

—, «The Theory of Games and the Evolution of Animal Conflict», *Journal of Theoretical Biology*, XLVII (1974), pp. 109-111.
—, *The Evolution of Sex*, Cambridge, Cambridge Unviersity Press, 1978.
—, «Hypercycles and the Origin of Life», *Nature*, CCLXXX (1979), pp. 445-446; reed, en Maynard Smith (1981:34-38).
—, «Symbolism and Chance», en J. Agassi y R. S. Cohen, eds. *Scientific Philosphy Today*, Hingham, Kluwer, 1981; reed, en Maynard Smith (1988:15-21).
—, *Evolution Now: A Century After Darwin*, San Francisco, Freeman, 1982.
—, «Adaptation and Satisficing», *Behavioral and Brain Sciences*, VI (1983), pp. 70-71; comentario sobre Dennett (1983).
—, «Structuralism Versus Selection. Is Darwinism Enough?», en Steven Rose y Lisa Appignanesi, eds., *Science and Beyond*, Oxford, Blackwell, 1986, pp. 39-46.
—, *Games, Sex and Evolution*, Londres, Harvester, 1988.
—, «What Can't the Computer Do?», *New York Review of Books*, 15 de marzo de 1990, pp. 21-25; reseña sobre Penrose (1989).
—, «Dinosaur Dilemmas», *New York Review of Books*, 15 de abril de 1991» pp. 5-7.
—, «Taking a Chance on Evolution», *New York Review of Books*, 14 de mayo de 1991, pp. 234-236.
Mayr, Ernst, «The Emergence of Evolutionary Novelties», en Sol Tax, ed., *Evolution after Darwin*, Chicago, University of Chicago Press, 1960, vol. I, pp. 349-380.
—, *The Growth of Biological Thought*, Cambridge, Harvard University Press, 1982.
—, «How to Carry Out the Adaptacionist Program», *American Naturalist*, CXXI (1983), pp. 324-334.
Mazlish, Bruce, *The Fourth Discontinuity: The Coevolution of Humans and Machines*, New Haven, Yale University Press, 1993; trad, esp.: *La cuarta discontinuidad: la coevolución de hombres y máquinas*, Madrid, Alianza, 1995.
McCulloch, W. S., y W. Pitts, «A Logical Calculus of the Ideas Immanent in Nervous Activity», *Bulletin of Mathematical Biophysics*, V (1943), pp. 115-133.
McGinn, Colin, *The Problem of Consciousness*, Oxford, Blackwell, 1991.
—, «In and Out of the Mind», *London Review of Books*, 2 de diciembre de 1993, pp. 30-31; reseña sobre Hilary Putnam, *Renewing Philosophy*.

McLaughlin, Brian, ed., *Dretske and His Critics*, Oxford, Blackwell, 1991.
McShea, Daniel W., «Arguments, Tests, and the Burgess Shale. A Commentary on the Debate», *Paleobiology*, IX (1993), pp. 339-402.
Medawar, Peter, «Unnatural Science», *New York Review of Books*, 3 de febrero de 1977, pp. 13-18.
—, *Pluto's Republic*, Oxford, Oxford University Press, 1982.
Metropolis, Nicholas, «The Age of Computing: A personal Memoir», *Daedalus* (invierno de 1992), pp. 119-130.
Midgley, Mary, «Gene-Juggling», *Philosophy*, LIV (1979), pp. 439-458.
—, «Selfish Genes and Social Darwinism», *Philosophy*, LVIII (1983), pp. 365-377.
Mill, John Stuart, *Utilitarianism*, *Fraser's Magazin*, 1861; reimp. 1863; trad, esp.: *El utilitarismo*, Madrid, Alianza, 1997
Miller, George A., «The Magical Number Seven, Plus or Minus Two», *Psychological Review*, LXIII (1956), pp. 81-97.
—, «A Very Personal History», Massachusetts Institute of Technology Cognitive Science Center, informe núm. 1, 1 de junio de 1979.
Millikan, Ruth, *Language, Thought and Other Biological Categories*, Cambridge, Massachusetts Institute of Technology Press, 1984.
—, *White Queen Psychology and Other Essays for Alice*, Cambridge, Massachusetts Institute of Technology Press, 1993.
Minsky, Marvin, «Why Intelligent Aliens Will Be Intelligible», en E. Regis, ed., *Extraterrestrials*, Cambridge, Cambridge University Press, 1985, pp. 117-128.
—, *The Society of Mind*, Nueva York, Simon &Schuster, 1985.
—, y Seymour Papert, *Perceptrons*, Cambridge, Massachusetts Institute of Technology Press, 1969.
Mivart, St. George, «Darwins Descent of Man», *Quarterly Review*, CXXXI (1871), pp. 47-90.
Monod, Jacques, *Le hasard et la nécessité*, Paris, Editions du Seuil, 1970; trad, esp.: *El azar y la necesidad*, Barcelona, Tusquets, 1989.
Moore, G. E., *Principita Ethica*, Cambridge, Cambridge University Press, 1903; trad, esp.: *Principia Ethica*, Barcelona, Laia, 1982.
Morgan, Elaine, *The Aquatic Ape*, Londres, Souvenir, 1982.
—, *The Scars of Evolution: What Our Bodies Tell Us About Human Origins*, Londres, Souvenir, 1990.
Muir, John, *Original Sanscrit Texts and the Origin and History of the People of India, Their Religion and Institutions*, Londres, Trubner, 1868-1873;

reed. Delhi, Oriental, 1972.
Murray, James D., *Mathematical Biology*, Nueva York, Springer-Verlag, 1989.
Nehamas, Alexander, «The Eternal Recurrence», *Philosophical Review*, LXXXIX (1980), pp. 331-356.
Neugebauer, Orto, «A Babylonian Lunar Ephemeris from Roman Egypt», en E. Leichtz, M. de J. Ellis y P. Gerardi, eds., *A Scientific Humanist: Studies in Elonor of Abraham Sachs*, Philadelphia, The Unviersity Museum, 1989, pp. 301-304.
Newell, Allen, y Herbert Simon, «The Logic Theory Machine», *IRE Transactions on Information Theory IT-2(3)* (1956), pp. 61-79.
—, «GPS: A Program That Simulates Human Thought», en Feigenbaum y Feldman [1964].
Newton, Isaac, *Philosophtae Naturalis Principia Mathematica*, Londres, Guil &John Innys, 172.6; trad, esp.: *Principios matemáticos de la filosofía natural*, Madrid, Alianza, 1998, 1 vols.
Nietzsche, Friedrich, *Morgenröte*, 1881; trad, esp.: *Aurora*, Madrid, Edaf, 1996.
—, *Die fröhliche Wissenschaft*, 1882; trad, esp.: *La gaya ciencia*, Barcelona, Alba, 1998.
—, *Jenseits von Gut und Böse*, 1885; trad, esp.: *Más allá del bien y del mal*, Madrid, Alianza, 1972.
—, *Zur Genealogie der Moral*, 1887; trad, esp.: *La genealogía de la moral*, Madrid, Alianza, 1972.
—, *Ecce Homo*, 1889; trad, esp.: *Ecce homo*, Madrid, Alianza, 1971.
—, *La voluntad de poderío (Der Willer zur Macht*, 1901).
Nowak, Martin, y Karl Simund, «A Strategy of Win-Stay, Lose-Shift Thar Outperforms Tir-for-Tar in the Prisoners Dilemma Game». *Nature*, CCCLXIV (1993), pp. 56-58.
Nozick, Robert, *Philosophical Explanation*, Cambridge, Belknap/Harvard University Press, 1981.
O'Neill, Onora, «The Perplexities of Famine Relief», en Tom Regan, ed., *Matters of Life and Death*, Nueva York, Random House, 1980, pp. 26-48.
—, *Faces of Hunger*, Boston, Allen Sc Unwin, 1986.
Orgel, Leslie E., y Francis Crick, «Selfish DNA: The Ultimate Parasite», *Nature*, CCLXXXIV (1980), pp. 604-607.
Otero, Carlos P, «The Emergence of "Homo Loquens" and the Laws of Physics», *Behavioral and Brain Sciences*, XIII (1990), pp. 747-750.

Pagels, Heinz, «A Cozy Cosmology», *The Sciences* (marzo-abril de 1985) pp. 34-39.
—, *The Dreams of Reason: The Computer and the Rise of the Sciences of Complexity,* Nueva York, Simon &Schuster, 1988; trad, esp.: *Los sueños de la razón: El ordenador y los nuevos horizontes de las ciencias de la complejidad*, Barcelona, Gedisa, 1990.
Paley, William, *Natural Theology: or, Evidences of the Existence and Attributes of the Deity, Collected from the Appearances of Nature*, Londres, Faulder, 1803, 5.ª ed.
Papert, Seymour, *Mindstorms: Children, Computers and Powerful Ideas*, Nueva York, Basic Books, 1980.
—, *The Children's Machine: Rethinking School in the Age of the Computer*, Nueva York, Basic Books, 1993; trad, esp.: *La máquina de los niños: replantearse la educación en la era de los ordenadores*, Barcelona, Paidós, 1995.
Papineau, David, *Reality ami Representation*, Oxford, Blackwell, 1987.
Parfit, Derek, *Reasons and Persons*, Oxford, Clarendon, 1984.
Parsons, William Barclay, *Engineers and Engineering in the Renaissance*, Cambridge, Massachusetts Institute of Technology Press, 1939; reimp. 1967.
Peacocke, Cristopher, *A Study of Concepts*, Cambridge, Massachusetts Institute of Technology Press, 1992.
Peckham, Morse, ed., *The Origin of Species by Charles Darwin: A Variorum Text*, Pittsburgh, University of Pennsylvania Press, 1959.
Penrose, Roger, *The Emperor's New Mind Concerning Computers, Minds, and the Laws of Physics*, Oxford, Oxford University Press, 1989.
—, «The Nonalgorithmic Mind», *Behavioral and Brain Sciences*, XIII (1990) pp. 692-705.
—, «Setting the Scene: The Claim and the Issues», Wolfson Lecture, 15 de enero de 1991.
—, «Nature's Biggest Secret», *Review of Books* (octubre de 1993)» pp. 78-82; reseña sobre Weinberg (1992).
Piatelli-Palmarini, Massimo, «Evolution, Selection, and Cognition: From "Learning" to Parameter Setting in Biology and the Study of Language», *Cognition*, XXXI (1989), pp. 1-44.
Pinker, Steven, *The Language Instinct*, Nueva York, Morrow, 1994; trad, esp.: *El instinto del lenguaje: cómo crea el lenguaje la mente*, Madrid, Alianza, 1996.

—, y Paul Bloom, «Natural Language and Natural Selection», *Behavioral and Brain Sciences*, XIII (1990), pp. 707-784.

Pittendrigh, Colin, «Adaptation, Natural Selection, and Behavior», en A. Roe y G. G. Simpson, eds., *Behavior and Evolution*, New Haven, Yale University Press, 1958, pp. 390-416.

Poe, Edgar Allan, «Maelzel's Chess-Player», *Southern Literary Messenger*, 1836; reed, en *Essays and Reviews*, Nueva York, Library of America, 1984, pp. 1253-1276; trad, esp.: «El jugador de ajedrez de Maelzel», en *Ensayos y críticas*, Madrid, Alianza, 1973.

—, artículo, *Southern Literary Messenger*, julio de 1836; reed. en J. M. Walker, ed., *Edgar Allan Poe: The Critical Heritage*, Nueva York, Routledge &Kegan Paul, 1986, pp. 89-90.

Popper, Karl y John Eccles, *The Self and Its Brain*, Berlín, Londres, Springer-Verlag, 1977; trad, esp.: *El yo y su cerebro*, Barcelona, Labor, 1985.

Poundstone, William, *The Recursive Universe: Cosmic Complexity and the Limits of Scientific Knowledge*, Nueva York, Morrow, 1985.

—, *Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb*, Nueva York, Doubleday Anchor, 1992.

Premack, David, Gavagai, *Or the Future history of the Animal Language Controversy*, Cambridge, Massachusetts Institute of Technology Press, 1986.

Putnam, Hilary, «Mind, Language and Reality», *Philosophical Papers*, Cambridge, Cambridge University Press, 1975, vol. II.

—, *The Faces of Realism*, LaSalle, Open Court, 1987; trad, esp.: *Las mil caras del realismo*, Barcelona, Paidós, 1994.

Quine, W. V. O., «On What There Is», en *From a Logical Point of View*, Cambridge, Harvard University Press, 1953, pp. 1-19.

—, *Word and Object*, Cambridge, Harvard University Press, 1960; trad, esp.: *Palabra y objeto*, Barcelona, Labor, 1968.

—, «Natural Kinds», en *Ontological Relativity*, Nueva York, Columbia University Press, 1969, pp. 114-138.

—, *Quiddities: An Intermittently Philosophical Dictionary*, Cambridge, Harvard University Press, 1987.

Rachels, James, *Created from Animals: The Moral Implications of Darwinism*, Oxford, Oxford University Press, 1991.

Rawls, John, *A Theory of Justice*, Cambridge, Harvard University Press, 1971; trad, esp.: *Teoría de la justicia*, Madrid, Fondo de Cultura Económica, 1979.

Ray, Thomas S., «An Approach to the Synthesis of Life», en C. G. Langron, C. Taylor, J. D. Farmer y S. Rasmussen, eds., *Artificial Life II*, Redwood City, Addison-Weslev, 1992, pp. 371-408.
Raymo, Chet, «Mysterious Sleep», *Boston Globe*, 19 de septiembre de 1988.
Raymond, Eric S., *The New Hacker's Dictionary*, Cambridge, Massachusetts Institute of Technology Press, 1993.
Rée, Paul, *Der Ursprung der moralischen Empfindungen*, Chemnitz, E. Schmeitzner, 1877.
Richard, Mark, *Propositional Attitudes*, Cambridge, Cambridge University Press, 1992.
Richards, Graham, «The Refutation That Never Was: The Reception of the Aquatic Ape Theory, 1972-1987», en Roede y otros (1991), pp. 115-126.
Richards, Robert J., *Darwin and the Emergence of Evolutionary Theories of Mind and Behavior*, Chicago, University of Chicago Press, 1987.
Ridley, Mark, *The Problems of Evolutions Oxford*, Oxford University Press, 1985.
—, *Evolution*, Boston, Blackwell, 1993.
Ridley, Matt, *The Red Quenn: Sex and the Evolution of Human Nature*, Nueva York, Macmillan, 1993.
Robb, Christina, «How and Why», *Boston Globe*, 5 de agosto de 1991, p. 38.
Robbins, Tom, *Even Cowgirls Get the Blues*, Nueva York, Bantam, 1976; trad, esp.: *También las vaqueras sienten melancolía*, Barcelona, Grijalbo, 1978.
Roede, Machteld, Jan Wind, John M. Patrick y Vernon Reynolds, eds., *The Aquatic Ape: Fact or Fiction*, Londres, Souvenir, 1991.
Rosenblatt, Frank, *Principles of Neurodynamics*, Nueva York, Spartan, 1962.
Rousseau, Jean-Jacques, *Discours sur l'origin et les fondements de l'inégalité parmi les hommes*, Amsterdam, Rev, 1755; tra. esp.: *Discurso sobre el origen de la desigualdad entre los hombres*, Barcelona, Alba, 1996.
Rumelhart, D., «The Architecture of Mind: A Connectionist Approach», en M. Posner, ed., *Foundations of Cognitive Science*, Cambridge, Massachusetts Institute of Technology Press, 1989, pp. 133-159.
Ruse, Michael, *Sociobiology: Sense or Nonsense?*, 2.ª ed. Dordrecht, Reidel, 1985; trad, esp.: *Sociobiologia*, Madrid, Cátedra, 1989.
—, ed., *Philosophy of Biology*, Londres, Macmillan, 1989.
—, y Edward O. Wilson, «The Evolution of Ethics», *New Scientist*, XVII (octubre de 1985), pp. 50-52; reed, en Ruse (1989).

Rushdie, Salman, «Born in Bombay», *London Review of Books*, 7 de abril de 1994, p. 4; véase también VV. AA., *For Rushdie*.

Russell, Bertrand, *A History of Western Philosophy*, Nueva York, Simon & Schuster, 1945; tra. esp.: *Historia de la filosofía occidental*, Madrid, Espasa-Calpe, 1994, 1 vols.

Ruthen, Russell, «Adapting to Complexity», *Scientific American* (enero de 1993), p. 138.

Samuel, A. L., «Some Studies in Machine Learning Using the Game of Checkers», *IBM Journal of Research and Development*, III (julio de 1959), pp. 211-229; reed, en Feigenbaum and Feldman [1964:71-105].

Schelling, Thomas, *The Strategy of Conflict*, Cambridge, Harvard University Press, 1960; trad, esp.: *Estrategia del conflicto*, Madrid, Tecnos, 1972.

Schiffer, Stephen, *Remnants of Meaning*, Cambridge, Massachusetts Institute of Technology Press, 1987.

Schopf, J. William, «Microfossils of the Early Archean Apex Chert: New Evidence of the Antiquity of Life», *Science*, CCLX (1993), pp. 640-646.

Schrödinger, Ernst, *What Is Life*, Cambridge, Cambridge University Press, 1967.

Schull, Jonathan, «Are Species Intelligent?», *Behavioral and Brain Sciences*, XIII (1990), pp. 63-108.

Searle, John, «Minds, Brains and Programs», *Behavioral and Brain Sciences*, III (1980), pp. 417-458.

—, *Minds, Brains and Science*, Cambridge, Harvard University Press, 1985; trad, esp.: *Mentes, cerebros y ciencia*, Madrid, Cátedra, 1985.

—, *The Rediscovery of the Mind*, Cambridge, Massachusetts Institute of Technology Press, 1991; trad, esp.: *El descubrimiento de la mente*, Barcelona, Crítica, 1996.

Sellars, Wilfrid, *Science, Perception and Reality*, Londres, Routledge Kegan Paul, 1963.

Selzer, Jack, ed., *Understanding Scientific Prose*, Madison, University of Wisconsin Press, 1993.

Shea, B. T., «Eskimo Cranofacial Morphology, Cold Stress and the Maxillary Sinus», *American journal of Physical Anthropology*, XLVII (1977), pp. 289-300.

Sherman, Paul W., Jennifer U. M. Jarvis y Richard D. Alexander, *The Biology of the Naked Mole-Rat*, Princeton, Princeton University Press, 1991.

Shields, W. M., y L. M. Shields, «Forcible Rape: An Evolutionary Perspective», *Ethology and Sociobiology*, IV (1983), pp. 115-136.

Sibley, C. G., y J. E. Ahlquist, «The Phylogeny of the Hominoid Primates, as Indicated by DNA-DNA Hybridization», *Journal of Molecular Evolution*, XX (1984), pp. 2-15.

Sigmund, Karl, *Games of Life: Explorations in Ecology; Evolution, and Behaviour*, Oxford, Oxford University Press, 1993.

Sills, David L., y Robert K. Merton, eds., *The Macmillan Book of Social Science Quotations*, Nueva York, Macmillan, 1991.

Simon, Herbert, *Models of Man*, Nueva York, Wiley, 1957.

—, y Allen Newell, «Heuristic Problem Solving: The Next Advance in Operations Research», *Operations Research*, VI (1958), pp. 1-10.

—, «Theories of Decision-Making in Economics and Behavioral Science», *American Economic Review*, XLIX (1959), pp. 253-283.

—, *The Sciences of the Artificial*, Cambridge, Massachusetts Institute of Technology Press, 1969; trad, esp.: *Ciencias de lo artificial*, Madrid, ATE, 1979.

—, y Craig Kaplan, «Foundations of Cognitive Science», en M. Posner, ed., *Foundations of Cognitive Science*, Cambridge, Massachusetts Institute of Technology Press, 1989» pp. 1-47.

Singer, Peter, *The Expanding Circle: Ethics and Sociobiology*, Oxford, Clarendon, 1981.

Skinner, B. F., *Science and Human Behavior*, Nueva York, Macmillan, 1953; trad, esp.: *Ciencia y conducta humana*, Barcelona, Martínez Roca, 1986.

—, *Verbal Behavior*, Nueva York, Appleton-Century-Crofts, 1957.

—, *Beyond Freedom and Dignity*, Nueva York, Knopf, 1971; trad, esp.: *Más allá de la libertad y la dignidad*, Madrid, Martínez Roca, 1986.

Skyrms, Brian, «Justice and Commitment», 1993.

—, «Sex and Justice», *Journal of Philosphy*, XCI (1994), pp. 305-320.

—, «Darwin Meets *The Logic of Decision*: Correlation in Evolutionary Game Theory», *Philosophy of Science* (diciembre de 1994), pp. 503-528.

Smolensky, Paul, «On the Proper Treatment of Connectionism», *Behavioral and Brain Sciences*, XI (1983), pp. 1-74.

Smolin, Lee, «Did the Universe Evolve?», *Classical and Quantum Gravity*, IX (1992), pp. 173-191.

Snow, C. P., *The Two Cultures, and a Second Look*, Cambridge, Cambridge University Press, 1963.

Sober, Elliot, «Holism, Individualism, and Units of Selection», *Philosophy of Science Association*, II (1981), pp. 93-121.

—, «The Evolution of Rationality», *Synthese*, XLVI (1981), pp. 95-120.

—, *The Nature of Selection: F. volutionary Theory in Philosophical Focus*, Cambridge, Massachusetts Institute of Technology Press, 1984.
—, ed., *Conceptual Issues in Evolutionary Biology*, Cambridge, Massachusetts Institute of Technology Press, 1984; 2.ª ed. 1994.
—, *Reconstructing the Past*, Cambridge, Massachusetts Institute of Technology Press, 1988.
Spencer, Herbert, *The Principles of Psychology*, 2.ª ed. London, Williams Norgate, 1870.
Sperber, Dan, «Anthropology and Psychology: Towards an Epidemiology of Representations», *Man*, XX (1985), pp. 73-89.
—, «The Epidemiology of Beliefs», en C. Fraser y G. Gaskell, eds., *The Social Psychological Study of Widespresad Beliefs*, Oxford, Clarendon, 1990, pp. 15-44.
—, «The Modularity of Thought and the Epidemiology of Representations», en Lawrence A. Hirschfeld y Susan A. Gelman, eds., *Mapping the Mind: Domain Specificity in Cognition and Culture*, Cambridge, Cambridge University Press, 1995.
Sperber, Dan, y Deirdre Wilson, *Relevance: A Theory of Communication*, Cambridge, Harvard University Press, 1986.
Stanley, Steven M., *The New Evolutionary Timetable: Tossils, Genes, and the Origin of Species*, Nueva York, Basic Books, 1981.
Sterelny, Kim, reseña sobre Dawkins (1986a), *Australasian Journal of Philosophy*, LXVI (1988), pp. 421-426.
—, y P. Kitcher, «The Return of the Gene», *Journal of Philosophy*, LXXXV (1988), pp. 339-360.
—, «Punctuated Equilibrium and Macroevolution», en P. Griffiths, ed., *Trees of Life*, Norwell, Kluwer, 1992, pp. 41-63.
—, reseña sobre Ereshefsky (1991), *Philosophical Books*, XXXV (1994), pp. 9-29.
Stetter, Karl O., R. Huber, E. Blochl, M. Kurr, R. D. Eden, M. Fielder, H. Cash e I. Vance, «Hypcrthermophilic Archaea Are Thriving in Deep North Sea and Alaskan Oil Reservoirs», *Nature*, CCCLXV (1993), p. 743.
Stewart, Ian, y Martin Golubitsky, *Fearful Symmetry: Is God a Geometer*, Oxford, Blackwell, 1992; trad, esp.: *¿Es Dios un geómetra?*Barcelona, Crítica, 1995.
Stove, David, «A New Religion», *Philosophy*, XXVII (1992), pp. 233-240.
Strawson, P. F., «Freedom and Resentment», *Proceedings of the British Academy*, XLVIII (1962), pp. 118-139.

Sturgeon, Nicholas, «Moral Judgment and Norms», *Ethics*, XLVI (1985), pp. 5-41.

Symons, Donald, «FLOAT: A New Paradigm for Human Evolution», en George M. Scherr, ed., *The Best of the Journal of Irreproducihle Results*, Nueva York, Workman, 1983.

—, «On the Use and Misuse of Darwinism in the Study of Human Behavior», en Barkow, Cosmides y Tooby (1992:137-62).

Tait, P. G., «Prof. Tait on the Formula of Evolution», *Nature*, XXIII (1880), pp. 80-82.

Teilhard de Chardin, Pierre, *The Phenomenon of Man*, Nueva York, Harper Brothers, 1959; trad, esp.: *El fenómeno humano*, Madrid, Taurus, 1986.

Thompson, D'Arcy W., *On Growth and Form*, Cambridge, Cambridge University Press, 1917.

Tooby, John, y Leda Cosmides, «The Psychological Foundations of Culture», en Barkow, Cosmides y Tooby (1992:19-136).

Trivers, Robert, «The Evolution of Reciprocal Altruism», *Quarterly Review of Biology*, IV (1971), pp. 35-57.

—, «Parental Investment and Sexual Selection», en B. Campbell, ed., *Sexual Selection and the Descent of Man*, Chicago, Aldine, 1972, pp. 136-179.

—, *Social Evolution*, Menlo Park, Benjamin-Cummings, 1985.

Tudge, Colin, «Taking the Pulse of Evolution», *New Scientist*, 24 de julio de 1993, pp. 32-36.

Turing, Alan, *ACE Reports of 1946 and Other Papers*, ed. B. E. Carpenter y R. W. Doran, Cambridge, Massachusetts Institute of Technology Press, 1946.

—, «Computing Machinery and Intelligence», *Mind*, LIX (1950), pp. 433-460.

—, «The Chemical Basis of Morphogenesis», *Philosophical Transactions of the Royal Society of London*, CCXXXVII (1952), pp. 37-72.

Turnbull, Colin, *The Mountain People*, Nueva York, Simon &Schuster, 1972.

Ulam, Stanislaw, *Adventures of a Mathematician*, Nueva York, Scribner's, 1976.

Unger, Peter, *Identity; Consciousness and Value*, Nueva York, Oxford University Press, 1990.

Uttley, A. M., *Information Transmission in the Nervous System*, Londres, Academic, 1979.

VV. AA., *For Rushdie*, Nueva York, Baziller, 1994.

Van Inwagen, Peter, *Metaphysics*, Oxford, Oxford University Press, 1993.

—, «Critical Study», *Nous*, XXVII (1993) pp. 373-339; reseña sobre Unger (1990).
Vermeij, Geerat J., *Evolution and Escalation*, Princeton, Princeton University Press, 1987.
Von Frisch, K., *A Biologist Remembers*, Oxford, Pergamon, 1967.
Von Neumann, John, *Theory of Self-reproducing Automata*, ed. póstuma Arthur Burks, Champaign-Urbana, University of Illinois Press, 1966.
—, y Oskar Morgenstern, *Theory of Games and Economic Behavior*, Princeton, Princeton University Press, 1944.
Vrba, Elizabeth, «Environment and Evolution: Alternative Causes of Temporal Distribution of Evolutionary Events», *Suid-Afrikaanse Tydskif Wetens*, LXXXI (1985), pp. 229-236.
Wang, Hao, *From Mathematics to Philosophy*, Londres, Routledge &Kegan Paul, 1974.
—, «On Physicalism and Algorithmism», *Philosophia Mathematical* I (1993), pp. 97-138.
Wason, Peter, «Regression in Reasoning», *British Journal of Psychology* LX (1969), pp. 471-480.
Waters, C. Kenneth, «Why the Anti reductionist Consensus Won't Sruvive the Case of Classical Mendelian Genetics», *Philosophy of Science Association*, I (1990), pp. 125-139; reed, en Sober [1994:401-417].
Wechkin, Stanley, Jules H. Masserman y William Terris, «Shock to a Conspecific as an Aversive Stimulus», *Psychonomic Science*, I (1964), pp. 47-48.
Weinberg, Steven, *Dreams of a Final Theory*, Nueva York, Pantheon, 1991; trad, esp.: *El sueño de una teoría final* Barcelona, Crítica, 1994.
Weismann, August, *The Germ Plasm: A Theory of Heredity*, Londres, Scott, 1893.
Wertheimer, Roger, «Philosophy on Humanity», en R. L., Perkins, ed., *Abortion: Pro and Con*, Cambridge, Schenkman, 1974.
Wheeler, John Archibald, «Beyond the End of Time», en Martin Rees, Remo Ruffini y John Archibald Wheeler, *Black Holes, Gravitational Waves and Cosmology: An Introduction to Current Research*, Nueva York, Gordon and Breach, 1974.
White, Stepehn L., «Self-Deception and Responsabilitv for the Self», en B. McLaughlin y A. Rorty, eds., *Perspectives on Self-Deception*, Berkeley, University of California Press, 1988.

—, *The Unity of the Self*, Cambridge, Massachusetts Institute of Technology Press, 1991.
—, «Constraints on an Evolutionary Explanation of Morality»; manuscrito inédito.
Whitfield, Philip, *From So Simple a Beginning: The Book of Evolution*, Nueva York, Macmillan, 1993.
Wilberforce, Samuel, «Is Mr Darwin a Christian?», *Quarterly Review*, CVIII (julio de 1860), pp. 225-264; reseña sobre *El origen de las especies*. publicada anónimamente.
Williams, Bernard, «A Critique of Utilitarianism», en J. J. C. Smart y Benard Williams, *Utilitarianism: For and Against*, Cambridge, Cambridge University Press, 1973, pp. 77-150.
—, «Evolution, Ethics, and the Representation Problem», en D. S. Bendall, ed., *Evolution from Molecules to Men*, Cambridge, Cambridge University Press, 1983, pp. 555-566.
—, *Ethics and the Limits of Philosophy*, Cambridge, Harvard University Press, 1985.
Williams, George C, *Adaptation and Natural Selection*, Princeton, Princeton University Press, 1966.
—, «A Defense of Reductionism in Evolutionary Biology», *Oxford Surveys in Evolutionary Biology*, II (1985), pp. 1-27.
—, «Huxley's Evolution and Ethics in Sociobiological Perspective», *Zygon*, XXIII (1988), pp. 383-407.
—, y Randolph Nesse, «The Dawn of Darwinian Medicine», *Quarterly Review of Biology*, LXVI (1991), pp. 1-22.
—, *Natural Selection: Domains, Levels, and Challenges*, Oxford, Oxford University Press, 1992.
Wilson, David Sloan, y Elliot Sober, «Reintroducing Group Selection to the Human Behavior Sciences», *Behavioral and Brain Sciences*, XVII [1994], pp. 585-608.
Wilson, E. O., *The Insect Societies*, Cambridge, Harvard University Press, 1971.
—, *Sociobiology: The New Synthesis*, Cambridge, Harvard University Press, 1975; trad, esp.: *Sociobiología*, Barcelona, Omega, 1980.
—, *On Human Nature*, Cambridge, Harvard University Press, 1978; trad, esp.: *Sobre la naturaleza humana*, Madrid, Fondo de Cultura Económica, 1983.

Wimsatt, William, «Reductionist Research Strategies and Their Biases in rhe Unir of Selection Controversy», en T. Nickles, ed., *Scientific Discovery: Case Studies*, Hingham, Kluwer, 1980, pp. 213-239.
—, «Units of Selection and the Structure of the Multi-Level Genome», en P. Asquith y R. Geire, eds., *Philosophy of Science Association,* II [1981], pp. 122-183.
—, «Developmental Constraints, Generative Entrenchment, and the Innate-acquired Distinction», en W. Bechtel, ed., *Integrating Scientific Disciplines*, Dordrecht, Martinus-Nijhoff, 1986, pp. 185-208.
—, y Monroe Beardsley, «The Intentional Fallacy», en *The Verbal Icon: Studies in the Meaning of Poetry*, Lexington, University of Kentucky Press, 1954.
Wittgenstein, Ludwig, *Tractatus Logico-philosophicus*, Londres, Routledge &C Kegan Paul, 1921; trad, esp.: *Tractatus logico-philosophicus*, Madrid, Alianza, 1997.
Wright, Robert, «The Intelligence Test», *New Republic*, 29 de enero de 1990, pp. 28-36; reseña sobre Gould (1989a).
Wright, Sewall, «Evolution in Mendelian Populations», *Genetics*, XVI (1931) p. 97.
—, «The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution», *Proceedings of the XI International Congress of Genetics*, I (1932), pp. 356-366.
—, «Comments on the Preliminary Working Papers of Eden and Waddington», en P. S. Moorehead y M. M. Kaplan, eds., *Mat-hematical Challenges to the Neo-Darwinian Interpretation of Evolution*, Filadelfia, Wistar Institute Press, 1967.
Young, J. Z., *A Model of the Brain*, Oxford, Clarendon, 1965.
Zahavi, A., «The Theory of Signal Selection and Some of Its Implications», en V. R Delfino, ed., *International Symposium on Biological Evolution*, Bari, Adriatici Editrici, 1987, pp. 305-327.



DANIEL C. DENNETT (Boston, 1942) es profesor universitario y catedrático Austin B. Fletcher de Filosofía, además de codirector del Centro de Estudios Cognitivos de la Universidad Tufts. Es miembro de la Academia Estadounidense de las Artes y las Ciencias y de la Academia Europea de Ciencias y Artes. Es uno de los filósofos de la ciencia más destacados en el ámbito de las ciencias cognitivas, especialmente en el estudio de la conciencia, la intencionalidad, la inteligencia artificial y de la memética. También son significativas sus aportaciones acerca de la significación actual del darwinismo y la religión. Ha escrito varios libros, incluyendo *Romper el hechizo*, *La evolución de la libertad*, *La peligrosa idea de Darwin* y, más recientemente, *Bombas de intuición y otras herramientas del pensamiento*.

Notas

[1] No voy a dedicar ningún espacio en este libro a hacer un listado de codas las imperfecciones del creacionismo, ni tampoco a defender mi condena del mismo. Considero que esta labor ya ha sido realizada de forma admirable por Kitcher [1982], Futuyma [1983], Gilkey [1985] y otros. <<

[2] Gilbert Ryle me contó esta típica muestra de hipérbole russelliana. A pesar de la distinguida carrera de Ryle como profesor Waynflete de filosofía en Oxford, él y Russell raras veces se encontraron, en gran medida, como me contó Ryle, a causa de la dominante posición de Russell en la filosofía académica después de la Segunda Guerra Mundial. En cierta ocasión, sin embargo, Ryle se encontró en el mismo compartimento con Russell durante un largo y aburrido viaje en tren, y tratando desesperadamente de entablar conversación con su mundialmente famoso compañero de viaje, le preguntó por qué creía que Locke, quien no era ni tan original ni tan buen escritor como Berkeley, Hume o Reid, había tenido mucha mayor influencia que éstos en el mundo de la filosofía de los países de lengua inglesa. Ésta fue su réplica y el comienzo de la única conversación provechosa que tuvo con Russell. <<

[3] La incapacidad de Descartes para pensar sobre el pensamiento como materia en movimiento es discutida extensamente en mi libro *La conciencia explicada* (1991). El libro de John Haugeland titulado *Artificial Intelligence: The Very Idea* (1985) es una admirable introducción a las vías filosóficas que hacen que pueda pensarse esta idea. <<

[4] William Paley desarrolló el argumento del diseño con bastante mayor detalle biológico en su libro *Teología natural* (1803) añadiendo muchos adornos ingeniosos. La influyente versión de Paley fue la inspiración real y el objetivo del rechazo de Darwin, pero el Cleantes de Hume capta toda la fuerza lógica y retórica del argumento. <<

[5] Gjertsen puntualiza que dos milenios antes Cicerón utilizó el mismo ejemplo con el mismo propósito: «Cuando tú ves un reloj de sol o de agua, sabes que te señala la hora debido a cómo está diseñado, no por azar. ¿Cómo puedes imaginar que el universo en su totalidad esté desprovisto de propósito y de inteligencia, cuando comprende todas las cosas, incluso estos artefactos y sus artífices?» (Gjertsen 1989:199). <<

[6] Esta situación se ha dado a menudo en la ciencia. Por ejemplo, durante años se disponía de muchas evidencias a favor de la hipótesis de la deriva de los continentes —es decir, que África y América del Sur habían sido adyacentes y que se habían separado—, pero hasta que fue concebido el mecanismo de las placas tectónicas, era difícil aceptar seriamente la hipótesis. <<

[7] Esta suma que apareció en la primera edición es errónea y cuando se le llamó la atención Darwin revisó sus cálculos para ediciones posteriores, aunque el principio general es irrefutable. <<

[8] Un ejemplo común de la regla de Malthus en acción es la rápida expansión de la población de levaduras introducida en la masa de pan fresco o en el mosto. Gracias al festín de azúcar y otros nutrientes, la explosión de población se mantiene durante varias horas en la masa del pan, unas pocas semanas en el mosto, pero pronto la población de levadura alcanza el techo malthusiano, debido a su propia voracidad y en la acumulación de productos de desecho: dióxido de carbono (el cual forma las burbujas que hacen crecer la masa del pan y el gas en el champán) y alcohol, los dos productos que nosotros, explotadores de las levaduras, tendemos a valorar. <<

[9] Como se ha señalado a menudo, Darwin no insistió en que la selección natural lo explicaba todo: era el «principal pero no el exclusivo sistema de modificación» (*ibidem*). <<

[10] Se ha sugerido a veces que la teoría de Darwin es sistemáticamente irrefutable (y en consecuencia científicamente necia) pero Darwin estaba al tanto de qué tipo de hallazgos debían producirse para refutar su teoría. «Aunque la naturaleza necesita de amplios períodos de tiempo para el trabajo de la selección natural, no requiere períodos indefinidos» (*ibidem*); así, si la evidencia geológica alcanzara a demostrar que no ha pasado bastante tiempo, su teoría total sería refutada. Esto deja aún una escapatoria, porque la teoría no fue formulable con suficiente rigor y detalle para decir justamente que muchos millones de años eran la mínima cantidad requerida, pero una escapatoria temporal que tiene sentido, dado que al menos algunas propuestas respecto a su tamaño podían ser evaluadas independientemente. (Kitcher [1985a:162-165] publicó una buena reflexión acerca de las posteriores sutilezas del argumento que impidió que la teoría de Darwin fuera directamente confirmada o rechazada). Otro famoso ejemplo: «Si se demuestra que existe cualquier organismo complejo que posiblemente no había sido formado por numerosas, sucesivas y ligeras modificaciones, mi teoría quedaría absolutamente destruida» (*ibidem*). Muchos han planteado este reto, pero como veremos en el capítulo ti, hay buenas razones para que no hayan tenido éxito en su demostración. <<

[11] El ideal de una ciencia deductiva (o «nomológico deductiva»), modelada sobre la física de Galileo o Newton, estaba completamente estandarizado, hasta muy recientemente, en la filosofía de la ciencia, de modo que no es sorprendente que se haya dedicado bastante esfuerzo a diseñar y criticar varias axiomatizaciones de la teoría de Darwin, dado que se presumía que entre tales formalizaciones se encontraba la vindicación científica. La idea, introducida en esta sección, de que Darwin debía ser considerado como postulando que la evolución es un proceso algorítmico, nos permite hacer justicia al innegable sabor *a priori* del pensamiento de Darwin, sin forzarlo en la inflexibilidad similar al lecho de Procrusto (y obsoleto) del modelo nomológico-deductivo. Véanse Sober [1984a] y Kitcher [1985a]. <<

[12] El 7 de abril de 1860. Véase Hardin (1964) para más detalles. <<

[13] Los informáticos a veces restringen el término *algoritmo* a programas de los que puede probarse que *finalizan*, que no tienen lazos infinitos, por ejemplo. Esta acepción específica, valiosa para algunos propósitos matemáticos, no es de mucha utilidad para nosotros. Por lo tanto, pocos de los programas de ordenadores en uso diario en todo el mundo se califican como algoritmos en este sentido restringido; la mayoría son diseñados para circular indefinidamente, esperando con paciencia las instrucciones (incluyendo las instrucciones para finalizar, sin las cuales continúan funcionando). Sus subrutinas son, sin embargo, algoritmos en sentido estricto, excepto cuando «virus» ocultos no detectados pueden hacer que el programa quede «colgado». <<

[14] Esta fascinante e incluso penosísima historia ha sido contada muchas veces pero aún provoca controversias. En primer lugar, ¿por qué retrasó Darwin la publicación de su libro? ¿Su comportamiento con Wallace fue generoso o monstruosamente injusto? Las difíciles relaciones entre Darwin y Wallace no son sólo un asunto de mala conciencia de Darwin respecto al modo de contestar a la correspondencia de Wallace en la que reclama su prioridad; como se comprueba aquí los dos estaban distanciados en su punto de vista y en su acritud respecto a la idea que ambos habían descubierto. Un buen tratamiento de esta cuestión se encuentra en Desmond y Moore [1991]; Richards [1987:159-161]. <<

[15] ¿De dónde procede el orden inicial? La mejor discusión la he encontrado en el libro de Penrose, *Cosmology tind the Arrow of Time*, 1989, cap. 7. **<<**

[16] Para revisiones accesibles de algunas de estas ideas véanse Pagels [1988], Stewart y Glubitsky [1992] y Langon y otros [1992]. <<

[17] ¡Nótese que no se deduce *lógicamente* que el ave esté relacionada con otras aves si encontramos su ADN casi idéntico en su secuencia con la de otras aves! «Tan sólo una coincidencia, no un plagio», sería una posibilidad lógica, pero que nadie tomaría en serio. <<

[18] Aunque no del todo imposible. Los satélites geoestacionarios, que orbitan al unísono de la rotación terrestre, son una especie de milagrosos ganchos celestes. Lo que los hace tan valiosos —lo que los convierte en inversiones financieras rentables— es que a menudo deseamos fijar algo (como una antena, una cámara o un telescopio) en un punto en el cielo. Los satélites no sirven para elevar, ya que deberían ser colocados muy altos. La idea ha sido estudiada cuidadosamente Por otra parte, un cable fabricado con la fibra artificial más fuerte debería tener cientos de metros de diámetro en la parte superior —podría ir disminuyendo hasta casi un invisible hilo de pescar— justamente para sostener su propio peso, aparte del que cargue. ¡Incluso si se pudiera fabricar tal cable, nadie desearía que cayese de su órbita en su propia ciudad! <<

[19] Schull (1990) es el responsable de la perspectiva que nos permite ver especies como capaces de «estar viendo» las mejoras de diseño, gracias a sus diferentes capacidades para la exploración fenotipica (comentarios en Dennett 1990a). <<

[20] La descripción hecha del efecto Baldwin por Robert Richards [1987:400-503] ha sido una de las mayores provocaciones y guías para mi pensamiento en este libro (véase más adelante mi discusión). Lo que me ha parecido particularmente valioso (véase mi revisión, Dennett 1989a) es que Richards no sólo comparte con Baldwin y muchos otros darwinianos un oculto anhelo por los ganchos celestes —o al menos una visceral insatisfacción con las teorías que insisten en las grúas—, sino que también tenía la honestidad intelectual y el valor para exponer y examinar su propia incomodidad con lo que está obligado a denominar «ultradarwinismo». El corazón de Richards está claramente con el de Baldwin, pero su mente no le permite fanfarronear o tratar de ocultar las grietas que observa en los diques que otros han tratado de levantar frente al ácido universal. <<

[21] Véase también su discusión de la peculiar versión del reduccionismo de Lewontin, Rose y Kamin (1984) —que Dawkins denomina con acierto su «duende privado»— en la segunda edición de *El gen egoísta*. <<

[22] ¡Sí!, encarnado» personificado. Acerca de esto pienso que lo que desearíamos decir sería reduccionismo *en espíritu*. <<

[23] Todo el mundo sabe cómo responder a esta cuestión retórica con otra: ¿Estás tan enamorado de la verdad a toda costa que desearías saber si tu amante te es infiel? Estamos en el punto de partida. Estoy a favor de una respuesta que diga: Yo amo tanto al mundo que estoy seguro de que deseo conocer toda la verdad acerca de él. <<

[24] Sin embargo, ha habido algunas importantes reuniones simbióticas de organismos pertenecientes a reinos diferentes. El gusano platelminto *Convouhtta roscoffensis* no tiene boca y no necesita comer dado que está lleno de algas que fotosintetizan su alimento (Margulis y Sagan 1986). <<

[25] Los cladistas (cuyos puntos de vista serán discutidos brevemente más tarde) son una escuela de taxonomistas que rechaza, por varias razones, el concepto de «especies progenituras» persistentes. Todo proceso que conduce a la formación de una especie tiene como resultado un par de especies hijas y la extinción de la especie progenitora, no importa en qué se asemeja una rama superviviente a su progenitora, comparada siempre con otra rama. **<<**

[26] La cuestión se complica aún más por la existencia de hibridación —en la que miembros de dos especies diferentes tienen una descendencia fértil—, un fenómeno que plantea cuestiones interesantes que están fuera de la línea que tratamos ahora. <<

[27] El epistemólogo evolucionista y psicólogo Donald Campbell ha sido quien ha desarrollado con más vigor las implicaciones de este aspecto del legado de Darwin. <<

[28] Obsérvese una importante diferencia entre los legados de la Eva Mitocondrial, y del Adán Cromosoma-Y: todos nosotros, hombres y mujeres, tenemos mitocondrias en nuestras células, pero todas proceden de nuestras madres; si eres varón tienes un cromosoma Y heredado de tu padre, si eres mujer —virtualmente todas, aunque no siempre— no tienes cromosomas Y. <<

[29] Los filósofos han discutido a menudo extraños ejemplos de individuos que nos son conocidos solamente a través de descripciones precisas, aunque generalmente han dedicado su atención a individuos tan aburridos como el espía más bajo (¿ha habido uno?). Yo sugiero que la Eva Mitocondrial es un ejemplo bastante más delicioso, sobre todo por ser de genuino interés teórico en la biología evolucionista. **<<**

[30] Hay, naturalmente, escritores que se pasan la vida escribiendo textos divertidos para las comedias televisivas, y los propios comediantes crean bastantes de su propia cosecha, pero, con escasas excepciones, éstos no son los creadores de los chistes («¿Sabes el de aquel tipo que…?») que corren de boca en boca. <<

[31] En 1982, François Jacob, biólogo galardonado con el premio Nobel, publicó un libro titulado *Lo posible y lo real* y me apresuré a leerlo esperando que fuera un ensayo esclarecedor de qué piensan los biólogos sobre las dificultades que plantea esta intrincada cuestión de la posibilidad. Para mi desilusión, el libro tenía muy poco que decir sobre este tema. Es un hermoso libro y con un gran título, pero en mi humilde opinión no se corresponden. El libro que me gustaría leer todavía no ha sido escrito, así que trataré de escribir yo mismo parte de él en este capítulo. <<

[32] Borges escogió cifras ligeramente diferentes: libros de 410 páginas con 40 líneas de 80 caracteres cada una. El número total de caracteres por libro es bastante cercano al mío (1.312.000 frente a 1.000.000), lo que no representa mucha diferencia. Yo escogí números redondos para un más fácil manejo. Borges eligió un conjunto de caracteres con solo 2$ miembros, bastante teniendo en cuenta las mayúsculas del español (con un espacio en blanco, una coma y un período como única puntuación), pero no las del inglés. Yo escogí la cifra más cómoda de 100 para tener espacio, sin ninguna duda, para las letras mayúsculas y minúsculas y puntuaciones de todas las lenguas de alfabeto romano. <<

[33] Stephen Hawking [1988:129] insiste en esta línea: «Hay algo así como diez millón millón millón millón millón millón millón millón millón millón millón millón millón (1 con ocho ceros después) partículas que podemos observar en la región del universo». Denton [1985] facilita una estimación de 1.070 átomos en el universo observable. Eigen [1992:10] calculó el volumen del universo como 1.084 centímetros cúbicos. <<

[34] La Biblioteca de Babel es finita pero, curiosamente, ¡contiene en sus estantes todas las oraciones gramaticales en inglés! Se trata de un número infinito y la biblioteca no lo es. Incluso cualquier oración en inglés, de cualquier longitud, puede ser dividida en fragmentos de 500 páginas, cada uno de los cuales se encuentra en algún lugar de la biblioteca. ¿Cómo es posible? Algunos libros pueden ser utilizados más de una vez. El caso más derrochador es el más fácil de comprender: dado que hay volúmenes en los que cada uno de ellos contiene un solo carácter y hay otros en blanco, el uso repetido de estos 100 volúmenes crearía cualquier texto de cualquier longitud. Como Quine puntualiza en su informativo y divertido ensayo «Universal Library» (en Quine 1987), si nos aprovechamos de la estrategia de reutilizar los volúmenes y traducir todo en el código ASCII que usa nuestro procesador de textos, podemos almacenar la totalidad de la Biblioteca de Babel en dos volúmenes extremadamente delgados, ¡en uno de los cuales está impreso un o y en otro aparece un 1! (Quine también recordó que el psicólogo Theodor Fechner propuso la fantasía de la biblioteca universal bastante antes que Borges). <<

[35] Quine (1987) acuñó el término «hiperastronómico» con el mismo propósito. <<

[36] *The Vale of Laughter* (1953), que sigue así: «Siéntete absolutamente libre. Llámame a cualquier hora del día o de la noche». De Vries pudo haber inventado el juego de ver cuán grande es el efecto (deletéreo o no) que se puede conseguir con un simple cambio tipográfico. <<

[37] La comparación de un genoma humano con los volúmenes de la galaxia de *Moby Dick* explica con facilidad algo que ocasionalmente confunde a la gente acerca del Proyecto del Genoma Humano. ¿Cómo pueden los científicos hablar de secuenciar (copiar) el genoma humano si cada genoma humano es diferente de todos los demás no sólo en uno sino en cientos de miles de lugares (*loci* en el lenguaje de la genérica)? Como los proverbiales copos de nieve o las huellas dactilares, no hay dos genomas humanos reales que sean iguales exactamente, incluso los de gemelos univitelinos (la probabilidad de errores que se desplazan de lugar siempre está presente, incluso en las células de un individuo singular). El ADN humano es fácilmente distinguible del ADN de otra especie cualquiera, incluso la del chimpancé, que es alrededor del 90 por ciento igual en cada locus. Todo genoma humano real que ha existido está contenido en una galaxia de posibles genomas humanos que se encuentra más que astronómicamente distante de las galaxias de los genomas de otras especies, todavía dentro de la galaxia hay mucho espacio para que dos genomas humanos no sean idénticos. Poseemos dos versiones de cada uno de nuestros genes, una procedente de nuestra madre y otra de nuestro padre. De estas versiones pasan a nosotros la mitad de sus propios genes, aleatoriamente seleccionados de aquellos que recibieron de sus padres, sus abuelos, pero dado que sus abuelos eran todos miembros del *Homo sapiens*, sus genomas son similares en casi rodos los *locis* así que no hay diferencia la mayoría de las veces que sus abuelos facilitan sus genes. Sus genomas, sin embargo, difieren en muchos miles de *loci* y en aquellas hendiduras, cuyos genes son conseguidos como un asunto de suerte; una moneda lanzada al aire en la maquinaria para formar la contribución de nuestros padres a nuestro ADN. Por otra parte, las mutaciones se acumulan al ritmo de 100 por genoma y por generación en los mamíferos. «Es decir, nuestros hijos tendrán cien diferencias respecto a nosotros y a nuestra esposa en sus genes, como resultado de los errores aleatorios de la copia realizada por nuestros enzimas o como resultado de mutaciones en los ovarios o los testículos a consecuencia de los rayos cósmicos» (Matt Ridley 1993:45). <<

[38] Esto es una hipersimplifieación que no considera el papel del mensajero ARN y otras complicaciones. <<

[39] Los realizadores cinematográficos nunca plantean el problema del *lector* del ADN. David Haig me ha llamado la atención sobre el hecho de que la elección de la rana por los realizadores cinematográficos pone de manifiesto un interesante error; una especie, sugiere Haig, de falacia de la Gran Cadena del Ser. «Los seres humanos, naturalmente, están más estrechamente relacionados con los dinosaurios de lo que éstos lo están con las ranas. El ADN humano habría sido mejor que el ADN de la rana. El ADN de las aves sería aún mejor». <<

[40] Un tema reciente que a menudo se trata entre los teóricos evolucionistas es que el «gencentrismo» más o menos aceptado actualmente ha ido demasiado lejos. De acuerdo con esta observación, la ortodoxia sobreestima en gran medida el alcance en que el ADN puede ser considerado como una receta, compuesta de genes, que especifica un fenotipo o un organismo. Quienes defienden esta tesis son los deconstruccionistas de la biología, que, al degradar el texto, elevan al lector al poder Es un tema útil para contrarrestar el gencentrismo hipersimplificado, pero una sobredosis resulta tan absurda como el deconstruccionismo en los estudios literarios. Esto se trata con más detalle en el capítulo 11. <<

[41] La idea de libertad de acción es algo que necesitamos presuponer en cualquier caso, siendo la mínima denegación del *actualismo*, la doctrina que sostiene que sólo lo real es posible. David Hume en su *Tratado de la naturaleza humana* (1739) hablaba de una «cierta holgura» que deseamos que exista en nuestro mundo. Esta es la holgura que previene a lo posible de encogerse apretadamente alrededor de lo real. Esta es la holgura que se presupone para *cualquier* uso de la palabra «puedo», una palabra de la que difícilmente podemos prescindir. Algunos han pensado que, si el *deterninismo* era verdad, el *actualismo* sería verdad —o dándole la vuelta, si el actualismo es falso, el indeterminismo debía de ser verdadero—, pero esto es muy dudoso. El argumento contra el determinismo sería desconcertantemente simple: este átomo de oxígeno tiene una valencia 2; por lo tanto, puede unirse con dos átomos de hidrógeno para formar una molécula de agua (*puede* en este mismo momento, lo haga o no lo haga), por lo tanto, algo que no es real es posible, así que el determinismo es falso. Hay argumentos impresionantes de la física que llevan a la conclusión de que el determinismo es falso, aunque éste no sea uno de ellos. Yo estoy preparado para asumir que el *actualismo* es falso (y que esta hipótesis es independiente de la cuestión determinismo/indeterminismo) aun si no puedo llegar a probarlo, aunque sea sólo porque la alternativa sería dejarlo todo e irse a jugar al golf o a otra cosa. Para una discusión más completa del *actualismo*, véase mi libro *Elbow Room* (1984), especialmente el capítulo 6: «¿Puedo haber hecho otra cosa?», del que he sacado el material de esta nota. Véase también la concurrente opinión de David Lewis [1986:cap. 17], acerca de un tema relacionado, como es la irrelevancia del problema del indeterminismo para nuestro sentimiento de que el futuro está «abierto». <<

[42] Otros han sacado utilidad del fenómeno QWERTY haciendo análisis similares: David [1985], Gould [1991a]. <<

[43] George Williams [1985:20] lo ha explicado de este modo: «He insistido alguna vez en que… “las leyes de la ciencia física más la selección natural pueden facilitar una completa explicación de cualquier fenómeno biológico” (Williams 1966:6-7). He adoptado una posición menos extrema y simplemente identifico la selección natural como la única teoría que un biólogo necesita además de aquellas de la física. Tanto la biología como la física necesitan contar con los legados históricos para explicar cualquier fenómeno del mundo real». <<

[44] Véase, por ejemplo, la discusión en Dawkins [1986:124-125], que concluye: «Por el contrario, una fuerte “presión de selección”, podemos solicitar clemencia por pensarlo, cabe esperar que conduzca a una rápida evolución. En su lugar, lo que encontramos es que la selección natural ejerce acción frenadora de la evolución. La línea base de la evolución, en ausencia de selección natural, es la máxima velocidad posible. Ésta es sinónimo de velocidad de mutación». <<

[45] ¿Los límites de la lógica pura son profundos o superficiales? Creo que algunos dependen de su obviedad. Una deliciosa parodia del pensamiento adaptacionista es la obra de Norman Ellestrand, titulada *Why are Juveniles Smaller Than Their Parents?* (1983), en la que se explora casi heroica mente, y conteniendo la risa, una serie de razones «estratégicas» para el «Pequeño Tamaño Juvenil» («Juvenile Small Size» o JSS). Concluye con una valiente mirada a las futuras investigaciones: «En particular, otra característica juvenil está aún más extendida que el JSS y merece una cuidadosa atención teórica, como es el hecho de que los jóvenes *siempre* parecen más jóvenes que sus padres». <<

[46] Seymour Papert [1993:90] describe la observación del «aprendizaje de un chico discapacitado» en una clase en la que contar con los dedos estaba prohibido: «Cuando se sentó pude ver lo impaciente que estaba por manipular los dedos. Pero se dio cuenta de que no podía ser así. Entonces lo observé mirando a su alrededor en busca de algo con lo que poder contar. No había nada a mano. Su frustración crecía. ¿Qué podía hacer? ¡Llegó la inspiración! Caminé hacia el chico y le dije quedamente: "¿qué te parecen los dientes?". Por la cara que puso me di cuenta de que había entendido mi propuesta y por la cara de la ayudante, que ella no lo entendió. ¡Enseño a los discapacitados, naturalmente!, me dije a mí mismo. Hizo sus sumas con una sonrisa contenida, obviamente divertido con aquella subversiva idea». (Cuando consideramos la utilización de lo primero que se tenga a mano como un posible movimiento forzado, merece la pena recordar que no todos los pueblos de nuestro planeta han utilizado el sistema decimal; los mayas, por ejemplo, usaban un sistema basado en 20 números). <<

[47] Los eruditos siguen trabajando. Con la ayuda de los ordenadores, los más recientes investigadores han demostrado «que el modelo del siglo XIX sobre la constitución y el descenso de nuestros manuscritos de Platón era tan hipersimplificado que debe considerarse erróneo. Este modelo, en su forma original, asume que todos los manuscritos existentes eran copias directas o indirectas de uno o más de los tres más antiguos manuscritos existentes, cada uno de ellos es una copia literal; las variantes en los manuscritos más recientes tenían entonces que ser explicadas ya por una corrupción del copista o una enmienda arbitraria, que se hacía acumulativa con cada nueva copia…» (Brumbaugh y Wells 1968:2; la introducción nos facilita un brillante cuadro del más reciente estado de la cuestión). <<

[48] Mi agradecimiento a Noel Swerdlow, que me contó esta historia durante la discusión que siguió a su conferencia «El origen de la teoría ptolomica de los planetas» en el Coloquio de Filosofía de la Tufts University, 1 de octubre de 1993, que posteriormente me facilitó el trabajo de Neugebauer y una explicación de sus puntos más interesantes. <<

[49] Los elementos genéticos transmitidos en la *Drosophila* son «parásitos intragenómicos» y probablemente tienen un efecto negativo en la adaptabilidad de sus organismos huéspedes, así que no debemos crearnos vanas esperanzas. Véase Engels [1991]. <<

[50] «Confieso que creo que la vacuidad del espacio fenotípico está llena de pistas falsas… Bajo la hipótesis nula de que no existen limitaciones, los senderos que se ramifican a través del espacio ocupado por este proceso constituyen un paseo aleatoriamente ramificado en un espacio multidimensional. La propiedad típica de tal paseo en un espacio multidimensional es que la mayor parte del espacio está vacío» (Kauffman 1995:19). <<

[51] Advertencia: los biólogos utilizan ya el término *macroevolución* en contraste con microevolución para referirse a los fenómenos evolutivos a gran escala; los modelos de especiación y extinción, por ejemplo, en contraste con el refinamiento de las alas o las modificaciones en la resistencia a las toxinas dentro de las especies. Lo que llamo la evolución de los macros no tiene casi nada que ver con la macroevolución en este sentido. Sin embargo, el término *macro* es tan apto para mis propósitos que he decidido mantenerlo y tratar de compensar sus insuficiencias con este remiendo (una táctica que la madre naturaleza utiliza a menudo). <<

[52] Precisamente por esta razón, Richard Dawkins presenta también una discusión y una elaboración de las ideas de Cairns-Smith en su libro *El relojero ciego*. Dado que la idea de Cairns-Smith y la elaboración de Dawkins son tan buenas lecturas para los no expertos, recomiendo que sean leídos por sus deliciosos detalles ya que nos aportan un sumario suficiente para despertar el apetito, añadiendo el consejo de que hay problemas con la hipótesis de Cairns-Smith y sopesando la advertencia con la confianza de que aun si la hipótesis fuera finalmente rechazada —una cuestión abierta— hay otras alternativas comprensibles con más dificultad para que puedan ser las siguientes en ser tomadas en serio. <<

[53] Küppers [1990:137-146] tomó prestado un ejemplo de Eigen [1976] para ilustrar la idea subyacente: un juego de «selección no darwiniana» que se puede jugar en un tablero de ajedrez con diferentes bolas de colores. Se empieza colocando aleatoriamente las bolas en todos los cuadros creando el efecto inicial de confeti. Ahora se lanzan dos dados (de ocho caras) para determinar el cuadrado (columna 5, fila 7, por ejemplo) en el que actuar. Se retira la bola de este cuadro. Se lanza de nuevo el dado; se va al cuadrado que ha salido, se comprueba su color y se coloca una bola de ese color justamente en el cuadro vacío («reproducción» de la bola). Se repite el proceso una y otra vez. Eventualmente, estas acciones tienen el efecto de desaleatorizar la distribución inicial de los colores, así que un color termina «ganando» pero no por razón alguna, sino, sencillamente, por suerte histórica. Küppers denomina a este proceso una «selección no darwiniana» porque la selección se hace en ausencia de una causa de sesgo; *selección sin adaptación* sería un término más familiar. No es darwiniana no sólo porque Darwin no vio la importancia de permitirlo, ni porque Darwin (o el darwinismo) no pudiera acomodarse a ella. Es evidente que puede. <<

[54] Esta descripción de la vida ha sido adaptada de una previa publicación mía (Dennett 1991b). Martin Gardner introdujo el juego de la vida en dos de sus columnas en *Scientific American*, en octubre de 1970 y febrero de 1971 sobre «Juegos Matemáticos»». Poundstone [1985] hace una excelente exploración del juego y de sus implicaciones filosóficas. <<

[55] Poundstone [1985] proporciona simulaciones simples en lenguaje ensamblador BASIC e IBM-PC, que uno puede copiar en su propio ordenador, y describe algunas interesantes variaciones. <<

[56] Véase Dennett [1987: cap. 9], para más implicaciones teóricas de este «asunto» entre espacio y tiempo. <<

[57] Para una perspectiva completamente diferente sobre la física bidimensional y la ingeniería, véase *The Planiverse* de A. K. Dewdney [1984], libro que supone una gran mejora respecto al *Flatland* de Abbott [1984]. <<

[58] John McCarthy ha explorado durante años la cuestión teórica de la mínima configuración del mundo de la vida que puede aprender la física de su propio mundo, y ha tratado de alistar a sus amigos y colegas en el estudio de esta cuestión. A mí me ha parecido siempre que la perspectiva de tal prueba es apetitosa pero los caminos para conseguirlo están muy alejados de mí. Hasta lo que conozco, nada sustantivo ha sido publicado hasta ahora sobre esta interesante cuestión epistemológica, pero deseo animar a que otros la sigan. El mismo experimento mental ha sido propuesto, independientemente, por Stewart y Golubitsky [1992:261-162]. <<

[59] Para un análisis más detallado de estos temas y una defensa del «neoplatonismo» de término medio, véase J. Leslie [1989]. (Como la mayoría de las posiciones intermedias, no se parece en nada a una posición escéptica, pero al menos es un ingenioso intento de compromiso). Van Inwagen (1993a: caps. 7 y 8) proporciona un claro e implacable análisis de los argumentos —del argumento de Leslie y también de otros argumentos que han sido presentados aquí— desde una posición de neutralidad inusual. Aquel que no quede del todo satisfecho con mi tratamiento del rema, debe volver a la fuente original. <<

[60] Para un atractivo examen de esta cuestión, véase el cap. 2 del libro de Robert Nozick *Philosophical Explanation*. Nozick ofrece como candidatas varias respuestas diferentes, todas ellas admite que son realmente extravagantes, aunque cautivadoras: «Sin embargo, la cuestión nos afecta tan profundamente que cualquier aproximación que renga una posibilidad de generar una respuesta parecerá extremadamente fantástica. Quienquiera que haya propuesto una respuesta que no sea extraña, demuestra que no entiende la cuestión» (Nozick 1981:116). **<<**

[61] Para una clara reconstrucción de la deducción de Nietzsche, inusualmente cuidadosa, que él describe como la «más científica de las hipótesis», consúltese a Danto 1965:201-209. Para una discusión y revisión y otras interpretaciones de Nietzsche sobre la notoria idea de la eterna recurrencia, véase Nehamas [1980], quien argumenta que para «los científicos Nietzsche significa específicamente lo no-teleológico». Un recurrente —pero no eterno— problema con la apreciación de la versión de Nietzsche del eterno retorno es que, al contrario que Wheeler, Nietzsche parece creer que ésta vida sucederá de nuevo no a causa de que ésta *y todas las posibles variaciones* sucederán una y otra vez, sino debido a que sólo hay una posible variación —ésta— y sucederá una y otra vez. Nietzsche, en resumen, parece creer en el actualismo. Me parece que esto no es esencial para una apreciación de las implicaciones morales que Nietzsche pensó que él podía o debía extraer de la idea y quizá para el saber de Nietzsche (pero ¿qué podemos hacer?). <<

[62] Descartes había planteado la cuestión de si Dios había creado las verdades de las matemáticas. Su seguidor Nicolás Malebranche (1683-1715) expresó con solidez que estas verdades no necesitan principio, siendo tan eternas como otras cosas podrían serlo. <<

[63] La similitud entre estos temas y los que he desarrollado en *La conciencia explicada* (1991) acerca de la necesidad de romper el Teatro Cartesiano con su «significador central», y distribuir su trabajo inteligente alrededor en una variedad de agentes periféricos, no es accidental. Sin embargo, este es, principalmente, un caso de evolución convergente, hasta donde puedo determinar. No había leído ninguno de los trabajos de Eigen cuando escribía mi libro, aunque seguro que me habrían inspirado si los hubiera leído. Un útil puente de unión entre los trabajos de Eigen sobre las moléculas y los míos sobre la conciencia, es el trabajo de Schull (1990) sobre la inteligencia de las especies y mi comentario, Dennett (1990a). <<

[64] Confío en que los filósofos reconocerán que Monod, al mismo tiempo, plantea y resuelve el problema de Putnam (1975) de la *Tierra Gemela*, al menos en el contexto del *problema de juguete* de la evolución molecular. El significado «no está en la cabeza» como Putnam observó, y tampoco está en el ADN. El problema de la *Tierra Gemela* conocido como el problema del contenido amplio frente al contenido estrecho, será desvelado brevemente en el capítulo 14, de modo que allí podré rendirle su apropiado funeral darwiniano. <<

[65] David Haig (comunicación personal) ha llamado mi atención sobre una fascinante estrategia en esta historia reveladora acerca de las proteínas plegadas: las acompañantes moleculares. «Las acompañantes son grúas moleculares *por excelencia*. Son proteínas con las cuales se asocia una cadena de aminoácidos mientras se pliegan, lo que le permite a la cadena adoptar una conformación que no sería posible en ausencia de la acompañante molecular. La acompañante es entonces desechada por la proteína ya plegada. Las acompañantes moleculares son estructuras evolutivamente muy conservadas… Las acompañantes moleculares han sido llamadas así porque cumplen funciones similares a las damas de compañía de las debutantes de un baile: su papel era estimular unas interacciones y dificultar otras». Para más recientes detalles, véase Martin y otros (1993), y Ellis y Van der Vies (1991). <<

[66] Eigen ha sugerido que hay una razón para que sean cuatro letras y no dos, aunque yo no voy a tomar una decisión sobre este punto. Quizás alguno de mis lectores podría calcular por sí mismo lo que esto significaría antes de ver lo que dice Eigen. Ya tiene en las puntas de sus dedos los principios relevantes de la ingeniería para dar con el resultado. <<

[67] Danny Hillis, el creador de la máquina de conexión, me contó una vez una historia sobre algunos *científicos* de la computación que diseñaron un componente electrónico para una aplicación militar (creo que formaba parte de un sistema de orientación en los aviones). Su prototipo tenía dos circuitos a bordo y el más elevado oscilaba, por lo que, buscando una fijación rápida, los científicos vieron un picaporte de latón en el laboratorio que justamente tenía el grosor apropiado. Lo extrajeron de la puerta y lo colocaron en el prototipo entre los dos circuitos de a bordo. Algún tiempo después uno de estos ingenieros fue llamado para examinar un problema que los militares tenían con los sistemas de fabricación y observó asombrado que entre los circuitos de a bordo en cada unidad existía un duplicado preciso de latón del picaporte original. Esta es una historia de Ur que tiene muchas variaciones bien conocidas en los círculos de ingeniería y entre los biólogos evolucionistas. Por ejemplo, véase a este respecto el divertido relato de Primo Levi acerca del misterio de los aditivos del barniz en *El sistema periódico* (1984). <<

[68] Consúltese Bedau [1991] para una exploración de este punto en la que llega a una conclusión algo diferente, y también a Unger [1990] con sus argumentos en contra. Unger insiste en que tenemos convenciones tales que deben estar —sobre bases lógicas— «a caballo» en tales circunstancias, de modo que una es el último ítem en Ja serie que falta *xy* el otro es el primero en la serie que posee *x* Pero, como ha observado, una conclusión más atractiva es: bastante peor para aquellas convenciones. <<

[69] Para algunos filósofos hay palabras bélicas. Como muestra de un claro intento de salvación de una lógica formal de las esencias que específicamente se dirige a los problemas planteados por la complejidad de los artefactos y organismos, véase Forties [1983, 1984]. La conclusión que saco del trabajo de Forbes es que consigue lo que puede ser una pírrica victoria sobre el firme escepticismo de Quine acerca de las esencias, pero en el proceso confirma su subyacente advertencia: contrariamente a lo que podamos creer no hay nada natural respecto al pensamiento esencialísta: ver el mundo a través de las gafas esencialistas no hace nuestra vida más fácil. <<

[70] Uno de los temas más importantes del filósofo alemán Martin Heidegger era que Sócrates debía ser culpado, en gran parte, de lo mal que le va a la filosofía, porque nos enseñó a pedir condiciones necesarias y suficientes. Dado que no es frecuente que Darwin y Heidegger se apoyen el uno al otro, merece la pena anotarlo en esta ocasión. Hubert Dreyfus ha mantenido durante mucho tiempo (por ejemplo, 1972 y 1979) que la inteligencia artificial está basada en un fallo en apreciar la crítica de Heidegger a Sócrates y piensa que aunque algunas partes de la IA pueden ser verdad, no es verdad en general, el cual está firmemente a favor de Darwin, una afirmación que defenderé más adelante en este capítulo, y con mayor detalle en los capítulos 13 a 15. <<

[71] Obsérvese aquí el paralelismo con mi discusión sobre la falsa dicotomía entre los modelos de la conciencia de Orwell y de Stalin en *La conciencia explicada*. En este caso tampoco existe marca *intrínseca* de lo canónico. <<

[72] De nuevo comprobamos la tolerancia de las imágenes cuando se utilizan sin orden ni concierto. Algunos teóricos hablan de *cuencas de atracción*, guiados por la metáfora de las bolas que ruedan ciegamente colina abajo hacia el *mínimum* local en lugar de escalar ciegamente colina arriba hasta el *máximum* local. Demos la vuelta a un *paisaje adaptativo* y las montañas se convertirán en cuencas, las crestas serán cañones y la «gravedad» correspondería al análogo de la presión de selección. No importa si escogemos «arriba» o «abajo» como la dirección favorecida. Aquí nos hemos deslizado, momentáneamente, a la perspectiva rival precisamente para insistir en este punto. <<

[73] El tema de la intencionalidad ha sido tratado extensamente por los filósofos de muy diferentes tradiciones en los últimos años. Para una revisión y una definición general, véase mi artículo «Intencionalidad» (con la colaboración de John Haugeland) en Gregory [1987]. Para un análisis más detallado, véanse mis primeros libros (Dennett 1969, 1978 y 1987b). <<

[74] Dos exposiciones muy accesibles del trabajo de Turing sobre la morfogénesis se encuentran en Hodges (1983:cap. 7) y Stewart y Golubitsky [1992]» los cuales también discuten su relación con los estudios teóricos más recientes en este campo de investigación. Aunque las ideas de Turing son interesantes, tienen como mucho una muy atenuada aplicación en los sistemas biológicos reales. John Maynard Smith (comunicación personal) recuerda haberse quedado fascinado por el trabajo de Turing (que le fue presentado por su supervisor J. B. S. Haldane) y durante años estuvo convencido de que «mis dedos deben ser ondas de Turing; mis vértebras deben ser ondas de Turing»; pero llegó a darse cuenta, resistiéndose a creerlo, de que todo esto no podía ser tan simple y tan bello. <<

[75] De hecho, el puente de unión entre ordenadores y evolución se encuentra aún más atrás en Charles Babbage, quien en 1834 concibió la «máquina diferencial» que se considera generalmente como la que inaugura la prehistoria de los ordenadores. ¡En su famoso libro *Ninth Bridgewater Treatise* (1813) aprovechó su modelo teórico de una máquina computadora para ofrecer una prueba matemática de que Dios tenía programada la naturaleza para generar las especies! «En la máquina inteligente de Babbage cualquier secuencia de números puede ser programada para interpolar, aunque sean largas, otras series que hayan estado funcionando. Por analogía. Dios había previsto en la Creación nuevos grupos de animales y plantas que aparecerían puntualmente a lo largo de la historia; él había creado las leyes que los producían, más que crearlos directamente» (Desmond y Moore 1991:213). Darwin conoció a Babbage y su *Treatise* e incluso asistió a sus *parties* en Londres. Desmond y Moore [1991:218] ofrecen algunas provocadoras visiones acerca del tráfico de ideas que podía haber cruzado este puente entre ambos.

Más de un siglo más tarde otra sociedad londinense de pensadores que amaban la mente, el Ratio Club, sirvió de vivero para ideas más recientes. Jonathan Miller llamó mi atención hacia el Ratio Club y me estimuló a investigar su historia mientras escribía este libro, pero no he averiguado casi nada hasta la fecha. Estoy sorprendido, sin embargo, con la fotografía de sus miembros en 1951 que adorna la portada del libro de A. M. Uttley, *Information Transmission in the Nervous System* (1979): Alan Turing está sentado en el césped, junto al neurobiólogo Horace Barlow (un descendiente directo de Darwin, dicho sea de paso); de pie, detrás, están Ross Ashby, Donald MacKay y otras figuras importantes de los inicios de lo que ha llegado a ser la ciencia cognitiva. ¡Qué pequeño es el mundo! <<

[76] La publicación de Samuel de 1959 fue reimpresa en la primera antología de la IA editada por Feigenbaum y Feldman, una obra clásica, *Computers and Thought* (1964). Aunque he leído este trabajo en Feigenbaum y Feldman cuando se publicó por primera vez, pasé por alto, como muchos lectores, muchos detalles y saboreé el momento más cargado de humor: una partida de 1969 entre el programa «adulto» y Robert Nealey, un campeón de damas. Nealey asumió la derrota con gracia: «En cuanto al juego final yo no me he enfrentado con tal competitividad de ningún ser humano desde 1954 cuando perdí mi ultima partida». Esta anécdota dio lugar a una soberbia conferencia de mi colega George Smith en un curso de introducción a las ciencias de la computación que codirigimos los dos en la Universidad de Tufts para reavivar mi interés sobre los detalles del artículo de Samuel en el que he encontrado algo nuevo y valioso cada vez que he vuelto a leerlo. <<

[77] Pronunciado para rimar con *stooge*, un *kludge* es un parche improvisado o *ad hoc*, o una reparación del software. Los puristas pronuncian esta palabra de jerga *kluge*, llamando la atención a su probable etimología en la deliberada e incorrecta pronunciación de la palabra alemana *klug(e)* que significa «listo»; pero de acuerdo con *The New Hacker's Dictionary* (Raymond 1993) el término puede tener un antepasado anterior, derivado del alimentador de papel Kluge, «un mecanismo adjunto a la prensa mecánica de imprenta» en uso en 1935. En su primera utilización designaba un «complejo e incomprensible artefacto con una función trivial». La mezcla de estima y desprecio que los *hackers* exhiben respecto a los *kluges* («¡cómo algo tan estúpido puede ser tan inteligente!») reproduce perfectamente la actitud de los biólogos cuando se maravillan ante las soluciones tan perversamente intrincadas que la madre naturaleza descubre a menudo. <<

[78] Este hecho ha sido aprovechado por los ingenieros contrarios a la ingeniería revertida. He discutido un ejemplo en Dennett [1978:279]:

«Hay un libro acerca de cómo detectar las antigüedades falsificadas (que también es, inevitablemente, un libro sobre cómo *hacer* antigüedades falsas) que ofrece este sagaz consejo para aquellos que desean burlarse del "experto" comprador: una vez que está terminada la mesa o cualquier otra cosa (habiendo utilizado todos los medios habituales para simular antigüedad y uso) cójase una taladradora eléctrica para hacer un agujero en la pieza en una situación llamativa pero incomprensible. El presunto comprador pensará: nadie practicaría este agujero que desfigura tanto la pieza sin alguna razón (esto no puede parecer "auténtico" de ninguna manera) así que lo debe haber hecho con algún propósito, lo que significa que esta mesa ha debido estar en uso en alguna casa: dado que ha sido usada en alguna casa, no ha sido hecha expresamente para su venta en la tienda de antigüedades… En consecuencia, es auténtica. Aunque esta "conclusión" da lugar a muchas dudas, el comprador estará tan preocupado imaginando usos para este agujero que tardará meses antes de que sus dudas vuelvan a aflorar».

Se ha afirmado, no sé con qué verosimilitud, que Bobby Fischer había utilizado la misma estrategia para derrotar a sus contrincantes en ajedrez, especialmente cuando el tiempo se estaba acabando: hacía deliberadamente un movimiento extraño y esperaba que el rival gastase un tiempo precioso en averiguar el sentido de dicho movimiento. <<

[79] «Sin embargo, esto no es explicar por qué, por ejemplo, hay vacuolas contráctiles en ciertos protozoos; es explicar por qué estos tipos de protozoos con vacuolas contráctiles ocurren» (Cummings 1975, en Sober 1984b: 394-395). <<

[80] Los tres párrafos precedentes han sido extraídos, con revisiones, de Dennett [1994a]. <<

[81] Guarro investigaciones clásicas de estos temas son *The Construction of Gothic Cathedrals* (1961), de John Fitchen, que se lee como una historia policíaca. *Building Construction Before Mechanization* (1986) de J. Fitchen, *Engineers and Engineering in the Renaissance* (1939, reed. MIT Press, 1967) de W. B. Parsons y *Engineers of the Renaissanc*e (1966) de Bertrand Gille. <<

[82] La evolución de la capacidad de evolucionar es para los darwinistas —retrospectivamente— un movimiento recursivo obvio —algo así como una fuente de «grúas»— y ha sido discutida por muchos pensadores. Para una discusión de este tema, véase Wimsatt [1981]. Para una diferente visión de este tema véase Dawkins [1989b]. <<

[83] Para un análisis más extenso de estos temas, véase Dennett [1990b]. <<

[84] Por si acaso, les aclaro que lo que hice fue imaginar que RxQ podría ser expresado en código Morse, y R en Morse es punto-raya-punto (el grupo de tres deslizadores cuenta como una raya). <<

[85] Yo introduje el término psicología popular en 1978 (Dennett 1981, 1987b) como denominación para el talento natural, quizá parcialmente innato, que los seres humanos tienen para adoptar posiciones intencionales. Véase Baron-Cohen [1995] y su fascinante contribución al estado actual del juego. Hay más acuerdos entre los filósofos y los psicólogos acerca de la existencia del talento que el que hay acerca de mi análisis de esta cuestión. Véanse las recientes antologías de este tema de Greenwood [1991] y Christensen y Turner [1993]. Véase Dennett [1987b], 1990b y 1991b para mi tratamiento de la cuestión. <<

[86] ¿No choca mi afirmación con las tesis de aquellos *cladistas* que proponen reducir la historia a un análisis estadístico de *caracteres* compartidos y no compartidos? (Para una revisión filosófica y una discusión, consúltese a Sober 1988). Sí, espero que sea y que mi revisión de sus argumentos (en gran parte a través de los análisis de Sober) me demuestre que las dificultades que ellos se crearon a sí mismos se deben en gran parte, si no enteramente, a sus intentos, tan esforzados, por encontrar vías no adaptacionistas para comprender las inferencias básicas que son claramente obvias para los adaptacionistas. Por ejemplo, aquellos cladistas que se abstienen de hablar de adaptación no pueden esperar ayuda por el obvio hecho de que tener un pie palmeado sea un «carácter» muy bueno y tener un pie sucio (cuando se examina) no lo es. Al igual que los conductistas que pretenden ser capaces de explicar y predecir «conductas» definidas en el lenguaje rigurosamente no interpretado de la trayectoria geográfica de las partes corporales, en lugar de utilizar el lenguaje funcionalmente rico de buscar, comer, ocultarse, cazar y así sucesivamente, los cladistas atemperados crean majestuosos edificios de intrincada teoría, los cuales son atractivos considerando que lo hacen con una mano atada tras su espalda, pero extraños, considerando que no habrían hecho esto si no hubiesen insistido en atarse una mano a la espalda (véase también Dawkins T986a:cap. 10, y Mark Ridley 1985:cap. 6). <<

[87] El mito de que el objetivo del trabajo de Gould y Lewontin era destruir el adaptacionismo y no corregir sus excesos, fue fomentado por la retórica de su escritura, aunque en algunos puntos se vuelve contra Gould y Lewontin, dado que los mismos adaptacionistas tendían a prestar mayor atención a la retórica que a los argumentos: «La crítica de Gould y Lewontin ha tenido poco impacto en la práctica, quizá porque han sido vistos como hostiles a la totalidad de la empresa y no simplemente como un tratamiento negligente de la cuestión» (Maynard Smith 1988:89). <<

[88] Kipling comenzó a publicar estas historias sueltas en 1897. <<

[89] Sin embargo, sir Alister Hardy, Linacre Professor de Zoología en Oxford, quien originalmente propuso la teoría, difícilmente pudo haber sido un miembro más firme del *establishment* científico. <<

[90] Por ejemplo, no se hace mención alguna de la teoría del mono acuático ni aun para rechazarla en dos recientes libros de mucha ilustración y poco texto, apropiados para la mesa del café, que incluyen capítulos sobre la evolución humana. Philip Whitfield en *From So Simple a Beginning: The Book of Evolution* (1993) ofrece unos cuantos párrafos sobre la teoría estándar del bipedalismo en la sabana. *The Primates' Progress*, de Peter Andrews y Christopher Stringer, es un ensayo bastante más largo sobre la evolución de los homínidos, en el libro *El libro de la vida* (Stephen Jay Gould, ed., 1993)». Pero éste, además, ignora la teoría del mono acuático —«the AAT»—. Y añadiendo insulto a olvido, ha habido también una desagradable parodia humorística de esta teoría realizada por Donald Symons [1983]» explorando la radical hipótesis de que nuestros ancestros volaron: «*The flying on air theory* o FLOAT, como es conocida por su acrónimo (con acrimonia entre el *establishment* reaccionario de la evolución humana)». Para una revisión de estas reacciones, véase G. Richards [1991]. <<

[91] El genetista Steve Jones [1993:20] nos ha planteado otra cuestión: existen más de trescientas especies llamativamente diferentes de peces cíclidos en el Lago Victoria. Entre ellos son muy diferentes; ¿cómo han llegado a serlo? «La opinión convencional es que el Lago Victoria en cierto momento se había secado y convenido en muchos lagos pequeños, lo que permitió que evolucionara cada especie. Aparte de los propios peces, no hay evidencia de que esto sucediera». Las historias adaptacionistas no han sido confirmadas y, en consecuencia, abandonadas. Mi ejemplo favorito es la ahora desacreditada explicación de por qué ciertas tortugas de mar emigran a través del Atlántico, entre África y Sudamérica, depositando los huevos en una orilla y alimentándose en la otra. De acuerdo con esta demasiado razonable historia, este hábito comenzó cuando África y Sudamérica comenzaron a separarse; en aquel entonces las tortugas atravesaban una bahía para desovar; la distancia creció imperceptiblemente a lo largo del tiempo, hasta que sus descendientes obligatoriamente cruzaron el océano para llegar a donde su instinto les decía que tenían que desovar. Resulta que el calendario en el que se produjo la ruptura de Gondwanaland (hipotético supercontinente del hemisferio sur que incluiría América del Sur, África, India, Australia y Antártida), es triste decirlo, no se acompasaba con el calendario de la evolución de las tortugas, aunque ¿no es ésta una idea atractiva? <<

[92] Recientemente Gould (1993a:318) ha descrito su posición antiadaptacionista como «el celo del converso» y en otro lugar confiesa: «a veces quisiera que todas las copias del libro *Ever Since Darwin* se autodestruyeran» (1991b:13), así que quizás hoy querría renunciar formalmente a estas palabras, lo que sería una pena, ya que expresan con elocuencia las razones del adaptacionismo. Sin embargo, la actitud de Gould respecto al adaptacionismo no es fácilmente discernible. *El libro de la vida* (1993) está repleto de razonamientos adaptacionistas que su lápiz rojo ha eludido y que, presumiblemente, cuentan con su apoyo. <<

[93] Dawkins no se contenta con que Sterelny descarte sus propias objeciones tachándolas de «evasivas» ya que, como señala (comunicación personal), éstas plantean un punto importante a menudo mal interpretado: «No es a los propios individuos humanos como Sterelny a quienes corresponde expresar su propio escepticismo de sentido común respecto a la proposición de que un 5 por ciento de aguijón es significativamente mejor que el 4 por ciento. Es una retórica fácil decir: “Veamos, ¿está usted realmente tratando de decirme que un 5 por ciento de aguijón realmente cuenta cuando se compara con un 4 por ciento?”. Esta retórica convence a menudo a los legos, pero los cálculos de la genética de poblaciones (como los realizados por Haldane) contradicen el sentido común de una manera fascinante y clarificante: debido a que la selección natural trabaja sobre genes distribuidos en muchos individuos y en una escala de muchos millones de años, la intuición humana actuarial está sobrepasada». <<

[94] Maynard Smith construyó su aplicación de la teoría de los juegos a la evolución sobre los fundamentos ya establecidos por R. A. Fisher [1930]. Una de las contribuciones más recientes de Maynard Smith ha sido demostrarle a Stuart Kauffman que él era, después de todo, un darwinista y no un antidarwinista (véase Lewin 1992:42,-43). <<

[95] A veces me pregunto si ha habido algún avance en el pensamiento de la segunda mitad de este siglo cuyo padre no haya sido Von Neumann. El ordenador, el modelo de autoreplicación, la teoría de los juegos y, por si fuera poco, Von Neumann ha hecho contribuciones mayores a la física cuántica. Independientemente de lo que por esto merece, yo sospecho que su formulación del problema de la medida en la mecánica cuántica es su idea fallida, una aceptación, mediante un juego de manos, de un modelo fundamentalmente cartesiano de observación consciente que ha confundido desde entonces a la mecánica cuántica. Mi alumno Turhan Canli abrió la puerta a esta cuestión (¡era un pregraduado!) en su tesis sobre el problema del gato de Schrödinger con la que desarrolló el esquema de una formulación alternativa de la física cuántica en la cual se cuantifica el tiempo. Si yo dominara la física (una posibilidad muy remota, triste es decirlo) seguiría la pista de este presentimiento, el cual podría extender por vías salvajemente ambiciosas mi teoría de la conciencia (Dennett 1991a); sin embargo, lo más probable es que yo siga siendo un espectador entusiasta, pero que comprende a medias, el desarrollo ulterior de esta cuestión. <<

[96] Para un fascinante relato de la historia de la teoría de los juegos y de sus relaciones con el desarmamento nuclear, véase el libro de William Poundstone *Prisoner's Dilemma: John von Neumann, Game Theory, and the Puzzle of the Bomb*, (1992). <<

[97] Como poderoso antídoto del paradigma de Pollyanna, véase G. Williams [1988]. <<

[98] El pesimista panglossiano dice: «¡Después de todo, no es una vergüenza que este sea el mejor de todos los mundos posibles!». Imaginemos el anuncio de una cerveza: cuando el sol se eleva sobre las montañas, uno de los cascarrabias que pierden el tiempo alrededor del fuego de un campamento declama: «No he conseguido una cerveza mejor que ésta», momento en el que su bella compañera rompe a llorar: «¡Oh no! ¿Realmente es verdad?». No se vendería mucha cerveza de esta marca. <<

[99] Gould está orgulloso de haber señalado el error de mirar atrás en el tiempo y ver linajes donde debemos ver «arbustos», incluyendo todos los fallos que no han dejado descendientes. Yo señalo un tipo contrario de error: imaginar densos (o incluso continuos) arbustos de posibilidades no realizadas donde de hecho puede haber escasos brotes abriendo caminos en avanzadillas relativamente aisladas dentro del enorme espacio de las posibilidades aparentes. <<

[100] «Sin embargo, cualquiera que sea el origen de la cúpula asentada sobre este tipo de pechinas (los *squinches*) la importancia de la cuestión ha sido, a mi juicio, enormemente exagerada. Las pechinas tipo *squinch* son un elemento de construcción que puede ser incorporado en casi todo tipo de arquitectura» (Krautheimer 1981:359). <<

[101] No soy el primero, como he descubierto recientemente, en destacar estos errores menores en la excursión de Gould por la historia del arte. Hace algunos años, dos biólogos evolucionistas se adelantaron a mis comentarios: Alastair Houston [1990] dirigió su atención al tema de las enjutas, las pechinas y los *squinches*, y Tim Clutton-Brock, en una conferencia pronunciada en Harvard, puso en duda la interpretación de Gould acerca de las bóvedas en abanico de la capilla del King’s College.

Es interesante notar que estas cuestiones pasaron inadvertidas para todos los deconstruccionistas y retóricos que han contribuido a los ensayos de un reciente libro [Selzer 199;] dedicado enteramente a un análisis de la retórica en el ensayo de Gould y Lewontin. Podemos suponer que alguno de los incluidos en este grupo de dieciséis humanistas se habría dado cuenta de los problemas reales en el dispositivo retórico fundamental del ensayo, aunque debe recordarse que estos sofisticados están interesados en el «conocimiento deconstructivo», ¡lo que significa que han trascendido la aburrida y pasada de moda dicotomía entre hecho y ficción y, por lo tanto, no están profesional mente interesados en si su lectura es la verdadera! <<

[102] En su propia discusión original del fenómeno QWERTY Gould [1991a:71] plantea una cuestión útil pero no la desarrolla posteriormente, hasta donde yo sé: debido a la curiosa secuencia de hechos que conducen a la adopción general de la disposición estándar QWERTY en las teclas de las máquinas de escribir: «Nunca se llevaron a cabo una serie de competiciones que hubiesen probado a QWERTY». Por lo tanto, es simplemente *irrelevante* preguntar si QWERTY es un diseño mejor que las alternativas X, Y y Z, dado que estas alternativas nunca se compararon con QWERTY en el mercado ni en el taller de diseño. Nunca aparecieron en un tiempo en el que, al parecer, podían haber marcado alguna diferencia. Los adaptacionistas deben estar alerta al hecho de que, aunque cualquier cosa que veamos en la naturaleza haya sido «probada contra todas las que se presenten a competir», sólo un subgrupo evanescentemente pequeño (y sesgado) de todas las competiciones imaginables tienen lugar. La inevitable parcialidad de todos los torneos significa que debemos ser cautelosos a la hora de caracterizar las virtudes de los ganadores. Un viejo chiste del Downeast recoge este sentido más sucintamente: «Buenos días, Edna». «Buenos días, Bessie, ¿cómo está tu marido?». «¿En comparación con qué?». <<

[103] «Durante los pasados treinta años, la teoría alopátrica (de la especiación) ha ganado en popularidad hasta llegar a ser, para la mayoría de los biólogos, *la* teoría de la especiación» (Eldredge y Gould 1972:92). Esta teoría ortodoxa tiene algunas llamativas implicaciones: «La teoría de la especiación alopátrica (o geográfica) sugiere una interpretación diferente de los datos paleontológicos. Si nuevas especies surgen muy rápidamente en poblaciones locales pequeñas periféricamente aisladas, entonces la gran expectativa de secuencias fósiles insensiblemente graduadas es una quimera» (Eldredge y Gould 1971:81). <<

[104] Para una introducción a este término véase el ensayo de Dietrich titulado «Macromutación» en el nuevo y excelente libro de consulta *Keywords in Evolutionary Biology* editado por Keller y Lloyd [1992]. <<

[105] George Williams [1992.130] discutió la importancia del estasis del traslado del hábitat, señalando que los parásitos, la «amplitud estacional de insolación» (cantidad de luz solar) y muchos otros factores ambientales, serán siempre diferentes después de un traslado geográfico, así que las poblaciones no serán nunca capaces de permanecer exactamente en el mismo ambiente selectivo y, en consecuencia, estarán sujetas a la presión de selección, a pesar del movimiento de traslado. Pero a mí me parece que gran parte, si no todo, del ajuste a estas presiones de selección puede permanecer invisible para el paleontólogo, el cual puede ver sólo en el registro de fósiles los cambios preservados en la parte dura del diseño. El traslado de hábitat puede ser responsable de muchos de los estasis observables paleontológicamente (¿qué otro estasis conocemos?) incluso si Williams está en lo cieno cuando dice que este estasis del diseño básico habrá enmascarado el concurrente «no-estasis» en la mayoría, si no todos, de los niveles en respuesta a los muchos cambios ambientales que deben acompañar a cualquier movimiento de traslado de hábitat de largo alcance. Y, a menos que muchas especies se muevan al unísono en sus traslados de hábitat, puede que no se trate de traslados de hábitat, dado que otras especies son elementos fundamentales para las especies en cualquier ambiente selectivo. <<

[106] Para una crítica similar de Gould, véase Ayala [1982]. Consúltese también G. Williams [1992:53-54]; Williams, que define la cladogénesis como el aislamiento, aunque de vida corta, de cualquier fondo de genes, también señala la trivialidad que para la teoría de la evolución tiene la cladogenesis a corto plazo (pp. 98-100). <<

[107] Las ideas de Gould acerca de la «separación de las especies a un nivel más alto» deben distinguirse de algunas ideas vecinas similares: las ideas acerca de la selección de grupos o selección de población, en la actualidad bajo intenso escrutinio y controversia entre los evolucionistas. Discutiremos estas ideas en el próximo capítulo. <<

[108] «Tinker to Evers to Chance» (Tinker pasa a Evers y Evers pasa a Chance) es un *meme* del béisbol, que inmortaliza la combinación de un doble juego de tres jugadores de la Sala de la Fama del béisbol, Joe Tinker, Johnny (el Cangrejo) Evers y Frank Chance, quienes jugaron juntos por Chicago en la Liga Nacional desde 1903 a 1912. En 1980, Richard Stem, un pregraduado en mi curso de introducción a la filosofía, escribió para mí un excelente ensayo, una puesta al día de los *Diálogos* de Hume; esta vez entre un darwiniano (Tinker, naturalmente) y un creyente en Dios (Evers, naturalmente), terminando, de modo apropiado, con Chance. Las múltiples convergencias sugeridas por este título —como buena suerte para hallar cosas valiosas por casualidad—, dado el enciclopédico conocimiento y amor por el béisbol del propio Gould, es sencillamente irresistible. **<<**

[109] Sí, ya sé que Joe Tinker jugaba entre la segunda y tercera base, no en la tercera base, pero, por favor, ¡permítaseme esta concesión! <<

[110] Gould escribe en respuesta a las objeciones de Conway Morris: «La explosión cámbrica fue demasiado grande, demasiado diferente y demasiado exclusiva» (1989a:230). Véanse también las anotaciones sobre la impredecibilidad de las trayectorias en «zigzag» (Gould 1989b). <<

[111] Philip Morrison ha puntualizado que si bien la proposición de que hay otra vida inteligente en el universo es mentalmente disparatada, también lo es su negación. No son verdades que funcionen en la cosmología. <<

[112] Por esta razón, precisamente, los biólogos han mostrado una mezcla de emociones acerca del reciente (aparente) descubrimiento hecho por J. William Schopf [1993] de microbios fósiles aproximadamente unos mil millones de años más antiguos (3,5 mil millones en lugar de 2,5 mil millones) de lo que la ortodoxia había supuesto. Si se confirmara, esto obligaría a revisar drásticamente muchas de las hipótesis estándar acerca de los límites intermediarios, dando más tiempo para la evolución de las formas avanzadas (¡vaya!) pero solamente para reducir el tiempo disponible que tendría el proceso de la evolución molecular para recorrer el camino que le separa de los microbios (¡Oh!). <<

[113] El método retórico de la «redescripción rimbombante» del lugar común fue primeramente descrito por Paul Edwards [1965] en un ensayo sobre otro oscurantista europeo, Paul Tillich. <<

[114] El libro de Teilhard tuvo un inverosímil defensor en Inglaterra, sir Julian Huxley, uno de los que contribuyeron —en realidad fue el que la bautizó— a la síntesis moderna del darwinismo. Como Medawar dejó claro, lo que Huxley admiraba en el libro de Teilhard era especialmente su apoyo a la doctrina de la continuidad de la evolución genética y «psicosocial». Esta es una doctrina que yo mismo he apoyado con entusiasmo bajo el título de unidad del espacio de diseño, de modo que algunos puntos de vista de Teilhard pueden ciertamente ser aplaudidos por algunos darwinistas ortodoxos. (Medawar pone reparos en este punto). Pero, en cualquier caso, Huxley no podía aceptar todo lo que Teilhard estaba ofreciendo. «Por todo ello, Huxley encontró que era imposible seguir a Teilhard “totalmente en su valiente intento de reconciliar los elementos sobrenaturales del cristianismo con los hechos e implicaciones de la evolución”. Pero ¡válgame Dios!, es precisamente esta reconciliación de lo que trata el libro de Teilhard» (Medawar 1981:251). <<

[115] Restrinjo el Lamarckismo a la herencia de las características adquiridas *a través del aparato genético*. Si relajamos la definición, entonces el lamarckismo no es claramente un falacia. Después de todo, los seres humanos heredan (por legados) la riqueza adquirida por sus padres y la mayoría de los animales heredan (por proximidad) los parásitos adquiridos de sus padres y algunos animales heredan (por sucesión) los nidos, madrigueras y guaridas adquiridos por sus padres. Todos estos son fenómenos de significado biológico, pero no son lo que Lamarck presentaba heréticamente. <<

[116] Dawkins [1986a:299] lanzó la advertencia correcta: el lamarckismo es «incompatible con la embriología tal como la conocemos», pero «esto no es decir que en algún lugar del universo, no puede haber algún sistema extraño de vida en el cual la embriología *sea* preformaciónista; una forma de vida que realmente tenga "un programa detallado de acción genético" y que realmente pueda, en consecuencia, heredar características adquiridas». Hay otras posibilidades que pueden ser denominadas también lamarckianas. Para una revisión de esta cuestión, véase Landaman [1991, 1993]; para otra interesante variación de este tema, véase el trabajo de Dawkins titulado «A Lamarkian Scare» ('Un sobresalto lamarckiano') (Dawkins 1982:164-178). <<

[117] Para aquellos que deseen estudiar estas y otras controversias de manera más completa, recomiendo los siguientes libros por ser particularmente claros y accesibles para los neófitos que *desean trabajar duro*: Buss [1987], Dawkins [1981], G. Williams [1992] y, con un manual muy valioso, Keler y Lloyd [1992]. Mark Ridley (1993) es un excelente libro de texto. Para un primer texto más accesible, Calvin [1986] es una excelente historia, con bastante especulación atrevida que estimula aún más el apetito. <<

[118] La fuente original, o al menos primaria, de *Cherchez la femme!* es la novela de Alejandro Dumas (Dumas padre) titulada *Los mohicanos de París* en la cual el inspector M. Jackal enunciaba el principio varias veces. Esta expresión ha sido atribuida a Talleyrand y a otros. (Mi agradecimiento a Justin Leiber por su indagación). <<

[119] Los amantes de los perros pueden protestar diciendo que hay buena evidencia de perros que sacrifican sus vidas por sus dueños humanos, colocando sus propias perspectivas de reproducción e incluso de «longevidad» personal firmemente en segundo lugar. Ciertamente esto sucede debido a que los perros han sido realmente *criados* para adquirir esta, en ocasiones fatal, lealtad a través de las especies. Sin embargo, se trata de casos excepcionales. El dibujante de comics, Al Capp comprendió este problema hace muchos años cuando creó sus deliciosos *shmoos*, unas globulosas masas blancas sin brazos con dos pseudópodos por pies y caras de gato con largos bigotes. Los *shmoos* aman a la gente por encima de todo e instantáneamente se sacrifican dondequiera que parezca apropiado, convirtiéndose a sí mismos en suntuosos bistecs (o sándwiches de mantequilla de cacahuete o cualquier otra cosa que sus compañeros humanos deseen). Los *shmoos*, podemos recordarlo, se reproducen asexuadamente por clonación en gran número en el acto; una licencia poética que consiguió Capp ante el persistente problema de cómo los *shmoos*, dadas sus proclividades, podían sobrevivir. ¡Kim Sterelny me ha sugerido que los *shmoos* exhiben el tipo de características que nosotros buscamos como una prueba de los intrusos intergalácticos en nuestro pasado! Si encontráramos organismos cuyas adaptaciones fueran manifiestas no para su propio directo beneficio, sino para el beneficio de sus hacedores putativos esto nos debería maravillar, aunque no debería ser concluyente. <<

[120] Véase la discusión en broma de Dennett [1991]. Algunos seres humanos afirman que les encanta dormir. «¿Qué planes tienes para este fin de semana?». «¡Dormir! ¡Ah, será maravilloso!». Otros seres humanos consideran esta actitud poco menos que incomprensible. La madre naturaleza no ve nada extraño en estas acritudes, bajo las condiciones correctas. <<

[121] Esto le ocurrió antes a un darwiniano tan esforzado como Thomas Henry Huxley, en su «Romanes Lecture» en Oxford (1893). «Los críticos de Huxley… tomaron nota de la aparente bifurcación que él había introducido en la naturaleza, entre los procesos naturales y la actividad humana, como si el hombre pudiera de alguna manera sacarse a sí mismo de la naturaleza» (Richards 1987:316). Huxley rápidamente se dio cuenta de su error e intentó restablecer un relato darwiniano de la cultura ¡mediante una llamada a la fuerza de la selección de grupo! ¡La historia se repite! <<

[122] Peter Kivy me informó, después de la Mandel Lecture, que este pasaje tan frecuentemente citado es falso y no corresponde a Mozart. Lo encontré en el clásico estudio de Jacques Hadamard *The Psychology of Inventing on the Mathematical Field* (1949:16) primeramente citado por mi mismo en Dennett [1975], una de mis primeras incursiones en el pensamiento darwiniano. Persisto en citarlo aquí, a pesar de la corrección de Kivy, porque no sólo expresa sino que ejemplifica la tesis de que los memes, una vez que existen, son independientes de autores y críticos. La precisión histórica es importante (por lo que he escrito esta nota al pie) pero el pasaje se adapta tan bien a mis propósitos que he decidido ignorar este pedigrí. Podía no haber persistido en esta idea, si no hubiera encontrado un meme de apoyo un día después de que Kivy me informara; yo escuchaba por casualidad a una guía en el Metropolitan Museum of Art, que comentaba el retrato de George Washington realizado por Gilbert Stuart: «Puede que no se parezca a George Washington pero así es como aparece ahora». Esta experiencia, naturalmente, ilustra otros de mis temas: el papel de la *serendipity* en toda obra de diseño. <<

[123] La confirmación de esta afirmación la dejo como un ejercicio para el lector. Entre los memes que estructuran la infoesfera y, en consecuencia, afectan a la transmisión de otros memes están las leyes de libelos. <<

[124] Para una buena discusión de la combatida relación entre lo que dice el gen y lo que dice la molécula, véase Waters [1990]. <<

[125] Generalmente la «acusación» de que la evolución cultural es lamarckiana es fruto de una confusión profunda, como Hull señala, pero en esta versión es innegable, aunque tampoco una «acusación». En concreto, la entidad que exhibe el talento lamarckiano de transmitir una característica adquirida no es el agente humano, sino el propio meme. <<

[126] Véase el punto paralelo sobre el grato —e incluso indispensable— poder que supone adoptar la posición intencional como una táctica científica en *heterofenomenologia*, la ciencia objetiva de la conciencia (Dennett 1991a). <<

[127] Para un análisis del impresionante poder mnemotécnico de la tradición oral, véase el clásico de Albert Lord ti rulado *The Singer of Tales* (1960) acerca de la tecnología de la memorización de versos desarrollada por los bardos desde los días de Homero hasta los tiempos modernos en los países balcánicos y en otras partes. <<

[128] Un ejemplo de este rechazo a Dawkins, lleno de vituperios y de incomprensión, por parte de un humanista que lo identifica como sociobiólogo se encuentra en Midgley [1979] un ataque tan desmesurado que no debería leerse sin su antídoto: Dawkins [1981]. Midgley [1983] es una réplica apologética pero todavía enormemente hostil. <<

[129] Dawkins [1989] ha presentado una importante y nueva perspectiva sobre los virus informáticos y su relación con otros memes. <<

[130] Los memes que son relativamente benignos para sus huéspedes pero perversos para otros, son frecuentes. Cuando el orgullo étnico se convierte en xenofobia, por ejemplo, es la imagen del fenómeno de un bacilo tolerable que por mutación se convierte en algo mortífero, si no necesariamente para su portador original, sí para otros. <<

[131] Lewontin, Rose y Kamin [1984:283] sostienen que estos memes presuponen una visión cartesiana de la mente, mientras de hecho los memes son un ingrediente clave (central pero opcional) en las mejores alternativas para los modelos cartesianos (Dennett 1991a). <<

[132] Esta sentencia apareció en el *Tufts Daily*, atribuida a Johann Wolfgang von Goethe, aunque supongo que es un meme de origen más reciente. **<<**

[133] Esta es una elaboración de ideas que he presentado primeramente en Dennett [1995]. Recientemente he descubierto que Konrad Lorenz [1973] describió una cascada similar de grúas, en términos diferentes, como es natural. <<

[134] Siendo imparciales con Chomsky, todo lo que él dice es que el libre albedrío *puede ser* un misterio. «No estoy propiciando esta conclusión sino simplemente haciendo notar que no puede descartarse *a priori*» (Chomsky 1975:157). ¡Esta moderada sugerencia ha sido impacientemente transformada por otros en una demostración con bases científicas! <<

[135] Otros dos libros de la biblioteca son las más apremiantes «refutaciones» de estas obras maestras, pero naturalmente la biblioteca no contiene ninguna refutación, propiamente dicha, de cualquiera de los libros *verdaderos*. Estos ataques maliciosos deben ser simplemente refutaciones *aparentes*; un ejemplo de un hecho que debe ser verdad, pero que carece sistemáticamente de utilidad, ya que no podríamos nunca decir qué libros estaban allí, sin la ayuda, digamos, de Dios. La existencia de este tipo de hechos llegará a ser importante en el capítulo 15. <<

[136] De hecho Chomsky ha revisado sus primeros puntos de vista acerca de la naturaleza del lenguaje, haciendo una distinción estos días entre el lenguaje E (el lenguaje externo —y se puede decir eterno— objeto platónico, inglés, en el cual muchos de los libros de la Biblioteca de Babel están escritos) y el lenguaje I (el lenguaje interno, intencional, idiolecto de un individuo), y niega que el lenguaje E sea un objeto apropiado para el estudio científico, así que él podría probablemente objetar la vía directa por la que he presentado esta objeción (en una comunicación personal a Steven Pinker). Pero existen más vías tortuosas para desarrollar este argumento y apelando únicamente al lenguaje I de los individuos. ¿Puede Chomsky, o cualquier otro, dar una buena razón para creer que cualquier libro de quinientas páginas, de frases cortas, que reúna los estándares del lenguaje I de cualquier individuo normal y culto, sería incomprensible («en principio») para esta persona? <<

[137] Fodor se ha enfrentado valientemente a esta dolorosa situación: «Nadie tiene la más ligera idea de cómo algo material puede ser consciente. Nadie se acerca ni siquiera remotamente a la más ligera idea de cómo algo material puede ser consciente» (Fodor 1992). En otras palabras, si pensamos que comprenderemos la cuestión de la conciencia, estamos equivocados. Tomemos sus palabras y cambiemos de tema, por favor. <<

[138] Este debate está recogido, con revisiones, en Dennett [1988a]. <<

[139] Cuando esto sucedió, Chomsky no pudo acudir y su lugar fue ocupado por su (y mío) buen amigo Massimo Piatelli-Palmarini (¡quien casi siempre está de acuerdo con Chomsky y raras veces conmigo!). Piatelli-Palmarini era el óptimo sustituto; había colaborado en la enseñanza de un seminario sobre cognición y evolución con Gould en Harvard y era el autor de un artículo (1989) que había hecho explícita por vez primera la posición de Gould y de Chomsky sobre el carácter no evolutivo del lenguaje. Su artículo fue una gran provocación y el objetivo del ensayo de Pinker y Bloom. <<

[140] El estilo confuso de Spencer fue blanco de la burla de William James en el epígrafe de la parte II (p. 233). Spencer [1870:396] había ofrecido la siguiente definición: «La evolución es una integración de materia y concomitante disipación de movimiento durante la cual la materia pasa de una indefinida incoherente homogeneidad a una definitiva y coherente heterogeneidad; y durante la cual el movimiento retenido sufre una transformación paralela». La memeología de la maravillosa parodia de James merece la pena ser recordada. Conseguí la cita de Garrett Hardin, quien me informó que él la había obtenido de Sills y Merton [1991:104]. Ellos a su vez citan las *Lecture Notes* de James [1880-1897] como su fuente, pero Hardin ha seguido la pista para obtener más detalles. P. G. Tail [1880:80] lo atribuye a un matemático llamado Kikman por su «exquisita traducción» de Spencer de la cual la versión de James —presumiblemente tomada prestada por Tait— es una mutación. Versión original de Kirkman (presumiblemente): «La evolución es un cambio desde una total similitud de la que no-se-puede-hablar en-absoluto, mediante continuas acciones-de-conjunción y acciones-de-algo-más». <<

[141] Este es también uno de los temas preferidos de Herbert Simon en *Ciencias de lo artificial* (1969), así que podemos llamarlo «simoniano» o «herbertiano». <<

[142] Uno de los alumnos de Warren McCulloch es Michael Arbib [1964], quien hizo importantes contribuciones a estos primeros desarrollos, y cuyas lúcidas discusiones sobre estos remas fueron para mí fuente de inspiración cuando era estudiante graduado, y cuyo último ensayo (1989) explora, en mi opinión, con persistencia nuevos territorios, en las trincheras y en otros suburbios, aún insuficientemente apreciados por muchos. <<

[143] Otros ensayos interesantes han sido los muchos modelos de aprendizaje, entendidos como «evolución darwiniana» en el sistema nervioso, realizados por neurocientíficos, recuperando el trabajo de Ross Ashby [1960] y J. Z. Young [1965] y continuando hoy en los trabajos de gente como Arbib y Grossber [1976], Changeux y Danchin [1976] y Calvin [1987], y también Edelman [1987] cuyo trabajo sería ciertamente un interesante ensayo si no se presentase como si fuera un salto en el desierto. <<

[144] El resto de esta sección se deriva de mi revisión del libro de Searle (Dennett 1993c). <<

[145] Dada la posición de Searle en este asunto, uno podría predecir que debe oponerse absolutamente a mi análisis del poder del pensamiento adaptacionista, como lo he presentado en el capítulo 9. Se opone. Yo no sé si él ha expresado este punto de vista en forma impresa, pero en varios debates conmigo (Rutgers, 1986; Buenos Aires, 1989) ha expresado el punto de vista de que mi tesis es exactamente un paso atrás: la idea de que uno puede cazar «las racionalidades que flotan libremente» de los procesos de selección evolutivos, es, desde su punto de vista, un travestismo del pensamiento darwiniano. Uno de nosotros se ha refutado a sí mismo de modo no intencionado; la identidad de la víctima la dejo como un ejercicio para el lector. <<

[146] Del mismo que los evolucionistas temen erróneamente que el efecto Baldwin cometiese el pecado del lamarckismo, y la precipitada huida del propio Darwin del catastrofismo, el rechazo indiscriminado de todo lo que huela a conductismo por los «mentalistas cabalmente modernos» (Fodor 1980) es un ejemplo de memes incorrectamente filtrados. Léase R. Richards [1987], autor de una excelente exposición de las distorsiones producidas por tales culpabilidades por asociación en el pensamiento evolucionista de los primeros tiempos, y Dennett [1975; 1978:cap. 4] por sus intentos de separar el trigo de la paja en el conductismo. <<

[147] Conviene leer a Fodor por diversión y para conocer en profundidad un rechazo tan profundo de la peligrosa idea de Darwin que desborda la práctica habitual de que, en principio, conviene acercarse con simpatía a la lectura de los textos. Su falsa exposición de las ideas de Millikan es particularmente llamativa, y no se debe confiar en ella en absoluto, aunque puede quedar compensada leyendo a la propia Millikan. **<<**

[148] Como suele suceder, estas cuestiones son más complicadas de lo que puedo presentar aquí (para más detalles, véanse Dennett 1987b, cap. 8, de donde se ha extraído este experimento, y Dennett 1990b, 1991c, 1991e, 1992). <<

[149] El libro *Dretske and His Critics* (McLaughlin, ed., 1991) se consagra ampliamente a este tema. <<

[150] Esto comenzó como una respuesta *impromptu* a la conferencia de Jaegwon Kim en Harvard, 29 de noviembre de 1990: «Emergence, Non-Reductive Materialism, and Downward Causation», y ha evolucionado ante los persistentes rechazos de Kim y muchos otros filósofos, a quienes les estoy agradecido. <<

[151] Para un ejemplo del mundo real de un proyecto de este tipo, véase el proyecto de la enorme enciclopedia CYC (abreviatura de «encyclopedia») en el MCC (Lenat y Guha 1990). La idea es codificar a mano los millones de hechos en una enciclopedia (más todos los otros millones de hechos que cada uno conoce, así que no hay posibilidad de colocarlos en la enciclopedia, como que las montañas son más grandes que las madrigueras de los topos y que las tostadoras no pueden volar) y entonces añadir una máquina de inferencias que pueda poner al día, conservar la coherencia, deducir sorprendentes implicaciones y, en general, servir a la base de conocimiento global. Para una aproximación completamente diferente a la IA, puede consultarse el proyecto del robot-humanoide de Rodney Brooks y Lynn Stein (Dennett 1994c). <<

[152] Algunos han argumentado que mi exposición de los patrones en Dennett [1991b] es epifenomenalista en cuanto al contenido. Esta es mi réplica. <<

[153] Dado que hay cajas de verdades, no se da apoyo por este medio a la hipótesis del «lenguaje del pensamiento» (Fodor 1975). Se supone que el conocimiento global fue almacenado en una forma casi lingüística precisamente para hacer el relato de la historia más fácil (¡ésta es probablemente la razón que ha motivado a la mayoría de los investigadores en ciencia cognitiva, que adoptan, fuera de toda conveniencia, la hipótesis del lenguaje del pensamiento!). <<

[154] Una primera versión de este experimento mental apareció en Dennett [1987b:cap. 8]. <<

[155] A la luz de este experimento mental, considere una cuestión planteada por Fred Dretske (comunicación personal) con admirable precisión: «Yo creo que nosotros podemos, de forma lógica, crear un artefacto que *adquiriese* intencionalidad original, pero no uno que (en el momento de la creación) la tuviese»». ¿Cuánto comercio con el mundo bastaría para transformar la escoria de la intencionalidad derivada en el oro de la intencionalidad original? Este es nuestro viejo problema del esencialismo, bajo un nuevo aspecto. Tiene el eco del deseo de mirar a través de un zoom un momento crucial donde de algún modo se identifica un umbral que marca el primer miembro de una especie, o el nacimiento de una función real, o el origen de la vida, y como tal manifiesta un fallo en la aceptación de la fundamental idea darwiniana de que todas estas excelencias emergen gradualmente a través de incrementos finitos. Nótese, además, que la doctrina de Dretske es una rama peculiar de un spencerismo extremo: el *actual* medio ambiente debe conformar el organismo antes de que la forma «cuente» como teniendo intencionalidad real; medios ambientes del *pasado*, filtrados a través de la sabiduría de ingenieros o de una historia de selección natural, no cuentan, aunque resultaran en las mismas estructuras funcionales. Hay algo erróneo y algo correcto en esto. Más importante que cualquier particular historia pasada de comercio individual apropiado con el mundo real es la disposición para implicarse en flexibles *futuras* interacciones, respuestas apropiadas a cualquier novedad que el mundo imponga. Pero —y esta es la sólida base, pienso, de la intuición de Dretske— ya que esta capacidad para el rápido rediseño es apta para mostrarse a sí misma en actuales o recientes patrones de interacción, su insistencia en que un artefacto exhibe una comprensión tipo «hágaselo usted mismo» (Dennett 1992) es verosímil, en tanto en cuanto desechemos el esencialismo y lo tratemos simplemente como un importante síntoma de intencionalidad merecedor de este nombre. <<

[156] El propio Maelzel es el inventor (o perfeccionador del metrónomo y fabricó la trompetilla para el oído de Beethoven, quien la utilizó durante años, desde que se quedó sordo. Maelzel también construyó una orquesta mecánica, llamada «Panharmonicon» para la cual Beethoven escribió la *Victoria de Wellington*, pero rompieron entre ellos a causa de los derechos de autor de dicha composición musical; Maelzel fue al mismo tiempo constructor de grúas y ladrón de grúas de gran talento. <<

[157] Un comentario hecho en 1971, citado en Wang [1993:133]. Véase también Wang [1974:316]: «Gödel cree que mecanismo en biología es un prejuicio de nuestro tiempo, el cual será refutado. En este caso, una refutación, en opinión de Gödel, consistiría en un teorema matemático sobre el hecho de que la formación dentro de los tiempos geológicos de un cuerpo humano por las leyes de la física (u otras leyes de similar naturaleza), comenzando desde una distribución aleatorizada de las partículas elementales y el campo, es tan improbable como la separación por el azar de la atmósfera en sus componentes». <<

[158] Uno que no comprende esto es Gerald Edelman, cuyas simulaciones de «darwinismo neural» son al mismo tiempo paralelas y fuertemente estocásticas (lo que implica aleatoriedad), un hecho que él cita, a menudo, erróneamente, como evidencia de que sus modelos no son algorítmicos, y que no está involucrado en la IA «fuerte» (Edelman 1992.). Sí lo está; por el contrario, sus protestas descubren una incomprensión elemental de los ordenadores, aunque precisamente demuestra, cómo todo el mundo sabe en IA, que mientras no se tenga la «absoluta ignorancia» (como Mackenzie escribió anónimamente; véase el capítulo 3, p. 97) no tendrá que comprender lo que está haciendo con el fin de hacer esto. <<

[159] Wechkin y otros [1964], Masserman y otros [1964]; para una discusión de este tema, véase Rachels [1991]. <<

[160] Este paralelismo fue quizá señalado por vez primera por E. G. Leigh: «Es como si tuviéramos que tratar con un parlamento de genes: cada uno actúa en su propio interés egoísta, pero si sus actos hieren a los otros, éstos combinarán sus esfuerzos para suprimirlo. Las reglas de transmisión de la meiosis desarrollan como una crecientemente inviolables reglas del juego limpio, una constitución diseñada para proteger al parlamento contra los actos nocivos de uno o de unos pocos. Sin embargo, a los loci tan estrechamente ligados a un distorsionador que los beneficios de Aprovecharse del éxito" sobrepasan el daño de su enfermedad, la selección tiende a potenciar el efecto de la distorsión. Así, unas especies deben tener muchos cromosomas si, cuando un distorsionador llega, la selección en la mayoría de los loci es para favorecer su supresión. Así como un parlamento tan pequeño puede ser pervertido por las intrigas de unos pocos, una especie con solo uno, estrechamente ligado cromosoma es una fácil presa para los distorsionadores» (Leigh 1971:249). Véase también Buss (1987:180ss), para una discusión de la secuestración de la línea germinal como básicamente una política de innovación que permite la vida multicelular. <<

[161] Rée era el amigo más querido de Nietzsche, con una amistad lo suficientemente estrecha como para que le confiara la tarea de transmitir a Lou Salomé la propuesta de matrimonio de Nietzsche en 1882, pero ella la rehusó, y Rée se enamoró de Lou. La vida es complicada. <<

[162] Es interesante observar que Nietzsche tenía una idea extensamente fundamentada y moderna acerca de las relaciones entre complejidad y cualquier noción del progreso global: «Las formas más ricas y más complejas —pues la expresión “tipo más elevado” no significa más que esto— perecen más fácilmente: sólo las formas más bajas mantienen una aparente indestructibilidad» (*La voluntad de poder*). **<<**

[163] «Para nosotros», en Asia, un individuo es una hormiga. Para usted él es un hijo de Dios. Este es un maravilloso concepto». (Lee Khan Yew, ministro *senior* de Singapur, en respuesta a las reclamaciones acerca de la sentencia de azotes de Michael Fay acusado de vandalismo, *Boston Globe*, 29 de abril de 1994, p. 8.) <<

[164] Según Wilson y Sober, los hutteritas tienen «la tasa de nacimientos más elevada de cualquier sociedad humana conocida», pero sería un error interpretar este dato como el triunfo del egoísmo reproductivo de Alexander. Sería un error táctico, por una razón: aunque hay muchos hutteritas o han sido, hay y ha habido muchas más monjas y monjas católicas, las historias de sus vidas serían muy difíciles de explicar como ejemplos de lucha individual, como siempre, por el campeonato reproductivo. Más expresivamente, si la idea del ejemplo de los hutteritas es destreza reproductiva del *grupo*, la tasa de nacimientos es relevante sólo cuando se utiliza como tasa de nacimiento de un grupo y no tenemos casi nada con qué comparar esta tasa dado que pocos grupos humanos, hasta donde yo sé, se comportan de esta manera. Quizá los hutteritas tienen tan elevada tasa de nacimiento individual porque muchos de sus hijos se van o son expulsados y han de ser sustituidos para mantener en funcionamiento las comunidades. ¡Podemos considerar la perspectiva verdaderamente maquiavélica de que es esto justamente lo que los egoístas genes desean, al fin y al cabo! Ellos encuentran un meme —el complejo hutterita— que sirvió sus propósitos y formó una facción: las espartanas comunidades hutteritas son realmente criaderos nada atractivos, de modo que muchos jóvenes las abandonarán, dejando sitio para más crías. No estoy apoyando esta versión, sino simplemente señalando que esta cuestión debería ser tratada como si se hubiera de dar una explicación evolucionista de cómo y por qué las comunidades hutteritas tienen estas características. <<

[165] Las complicaciones abundan, como es habitual. En algunas especies de escarabajos, los machos hacen una gran inversión en tapones de alimentos (con esperma adherido) por las que las hembras compiten. Este es un tipo de inversión de los padres, pero no del tipo que estamos discutiendo aquí. <<

[166] Gould comentó la misma llamativa estadística en «A Thousand Acts oí Kindness»» (‘Mil actos amables’) en Gould [1993d]. <<

[167] Si desea conocer las probabilidades de ligar una escalera de color en el póquer, un método es solucionar la ecuación proporcionada por la teoría de los juegos; obtendrá una respuesta definitiva. Otro método es jugar unos cuantos miles de millones de manos de póquer, barajar bien entre cada una de ellas, y simplemente contar las escaleras de color y dividir por el total de las manos jugadas. Este método proporciona una estimación *muy* fiable pero no es oficialmente infalible. El último método es la únicamente factible vía para estudiar los complicados escenarios de la ética evolucionista, pero, como ya hemos visto en la discusión (en el capítulo 7) de la reacción de Conway a las vías en las que ha sido explorado su juego de la vida, los resultados de tales simulaciones pueden ser engañosos y deben ser tomados a menudo con un grano de sal. <<

[168] Philip Kitcher inicia su revisión crítica de la sociobiología. *Vaulting Ambition* (1985), con una historia que no tiene respuesta acerca del daño infringido por el conocido examen británico once-plus —ahora abandonado gracias a Dios—, el cual marcaba a los niños de once años con un veredicto de arriba o abajo sobre lo que prometían, el cual fijaba inexorablemente el conjunto de caminos que sus vidas pudieran tomar. **<<**

[169] Un útil ejercicio cuando se consideran tales casos es imaginarse la creación de una habitación llena de robots toscamente racionales (inteligentes, pero sin herencia genética, en todo caso) y preguntarnos si llegarían a adquirir la conducta en cuestión. (Si el caso es complejo, una simulación en el ordenador debe ser utilizada, como una prótesis para nuestra imaginación). Si es así, no es tan sorprendente que los seres humanos hagan esto por todas partes, y esto probablemente no tiene nada que ver con su herencia de los primares, de los mamíferos e incluso de los vertebrados. <<

[170] Incluso Donald Symons [1992:142] sucumbió a este atrayente eslogan: «No hay tal cosa como un “solucionador de problemas general” porque no hay tal cosa como un problema general». ¿Oh? No hay tal cosa como una herida general; cada herida tiene una forma completamente característica, pero puede existir un sistema de curación general de las heridas, capaz de curar heridas de una variedad casi ilimitada de formas, simplemente porque es más barato para la madre naturaleza crear un sistema de curación general de las heridas que un sistema de curación especializado de las heridas (G. Williams 1966:86-87; véase también Sober 1981b:106ss.). Hasta qué punto es general cualquier mecanismo cognitivo, o hasta qué punto puede hacerse a través de una potenciación cultural, siempre será una cuestión empírica abierta. <<

[171] El material de este capítulo ha sido tomado de Dennett (1988b), donde estos temas se han desarrollado con más detalles. <<

[172] Merece la pena recordar que las matemáticas y la física son las mismas en todo el universo y que, en principio, podrían ser descubiertas por alienígenas (si los hubiera), no importa su clase social, predilecciones políticas, sexo (¡si tuvieran sexo!) o preferencias. Menciono esto para protegerme de la reciente estupidez que pueden haber escuchado originada en algunas de las escuelas de pensamiento —hablo libremente— en la sociología de la ciencia. Es descorazonador leer a un pensador tan sabio como John Patrick Diggins, cautivado por esta explicación:

Pero, como el señor Marsden señala, en el pasado se asumía que la ciencia debería ser el árbitro de tales disputas, mientras que hoy el papel de la ciencia se limita a ser otra simple vía para describir el mundo verbalmente, más que para conocerlo filosóficamente. En el pasado reciente, la religión fue eliminada del campus universitario porque le faltaban las credenciales científicas. Pero dado que este criterio ha perdido sus propias credenciales, el señor Marsden se pregunta por qué la religión no puede reclamar su lugar en los campus universitarios. Es correcto plantear esta cuestión (Diggins 1994).

No es «cientificismo» conceder objetividad y precisión a la buena ciencia, del mismo modo que no es adoración de la historia conceder que Napoleón durante un tiempo dominó a Francia y que el Holocausto sucedió realmente. Aquellos que temen los hechos tratarán siempre de desacreditar a los que los encuentran. <<

[173] Probablemente ha sido Axelrod (1984) con su derivación del golpe por golpe quien ha estado mis cerca de conseguir un «resultado» en este campo, pero, como él mismo puntualiza, las virtudes probables de la regla asumen condiciones que solamente se cumplen intermitentemente y, además, son controvertidas. En especial la denominada «sombra del futuro» debe ser «suficientemente grande», una condición con la cual, según parece, la gente razonable puede estar en desacuerdo indefinidamente. <<

[174] ¿Cómo iba a ser una buena cosa lo ocurrido en la central nuclear de Three Mile island en Estados Unidos? Al haber estado *cerca* de la catástrofe se convirtió en la alarma que nos alejó de lo que podrían haber sido desventuras del calibre de la de Chernobyl, por ejemplo. Seguramente mucha gente estaba *esperando* fervientemente que sucediera tal cosa y podía haber dado los pasos necesarios para que sucediera. El mismo razonamiento moral que condujo a Jane Fonda a hacer la película *El síndrome de China* (la ficción de una casi catástrofe en una planta nuclear) podía haber llevado a alguien en una situación distinta» a provocar el problema de Three Mile Island. <<

[175] Judith Jarvis Thomson ha objetado (en un comentario sobre «The Moral First Aid Manual» en Ann Arbor, 8 de noviembre de 1986) que ni «compra barato y vende caro» ni su contrapartida «consecuencialista», «haz más bien que mal», son estrictamente expresiones vacuas; ambas presuponen algo acerca de los fines últimos, dado que el primer consejo sería un mal consejo para aquel que pensó perder dinero, y el segundo no sería el interés último de toda la gente con una conciencia moral. Estoy de acuerdo. La segunda expresión compete, por ejemplo, al consejo que el rey pirara da a Federico, el peculiar «esclavo del deber» en *Pirates of Penzance*: «¡Amigo mío! Siempre cumple con tu deber y que tengas fortuna con las consecuencias». Ningún eslogan es *completamente* vacuo. <<

[176] Una kantiana que acusa al utilitarismo con la acusación de imponderabilidad práctica, con especial fuerza y claridad, es Onora O'Neill [1980]. Ella demuestra cómo dos utilitaristas, Garrett Hardin y Peter Singer, armados con la misma información, llegan a consejos opuestos sobre el acuciante dilema moral del alivio del hambre: debemos poner en marcha medidas drásticas para prevenir los esfuerzos miópicos para alimentar a las víctimas del hambre (Hardin) o debemos poner en marcha medidas drásticas para facilitar alimentos a las víctimas actuales del hambre (Singer). Para una consideración más detallada, véase O'Neill [1986]. Un crítico independiente es Bernard Williams, quien sostiene que el utilitarismo hace «enormes demandas sobre información supuestamente empírica, acerca de las preferencias de los pueblos, y esta información no sólo no está disponible en su mayor parre, sino que oculta una dificultad conceptual; pero esto es contemplado a la luz de una dificultad técnica o práctica, y el utilitarismo apela a una forma de pensar en la cual la dificultad técnica, incluso la insuperable dificultad técnica, es preferible a la falta de claridad moral, sin duda porque es menos alarmante. (Esta forma de pensar es de hecho profundamente estúpida…)» (1973:137). <<

[177] La sugerencia de un orden temporal en las cinco tases no es esencial, naturalmente. La poda arbitraria del árbol explorado con una búsqueda aleatoria, la puesta en marcha de la decisión por una evaluación parcial y no óptima de los resultados y la supresión de la necesidad de la segunda conjetura, no sigue la secuencia en el tiempo que yo diseñé en el ejemplo inicial. El proceso a este nivel es lo que he descrito en el modelo de múltiples diseños de la consciencia humana, en Dennett (1991a). <<

[178] Stephen White [1988] discutió el bien conocido intento de Strawson (1962) de terminar la exigencia de una justificación de «nuestras actitudes reactivas» mediante un dato en bruto sobre de nuestro modo de vida, acerca del cual «no tenemos elección». Demostró que el moderador no pudo resistir una posterior demanda de justificación (lo que White aporta en una ingeniosa vía indirecta). Véase también White (1991). Para una complementaria y clarificadora aproximación al problema práctico de la toma de decisiones ética, véase Gert (1973). <<

[179] Robert Axelrod me ha puntualizado que lo que Hofstadter llama el «dilema del lobo» es formalmente idéntico a la parábola de la «caza del venado» de Jean-Jacques Rousseau, en el *Discurso sobre et origen y los fundamentos de la desigualdad entre los hombres* (1755). Para una ulterior discusión de anticipaciones y dificultades, véase Dennett [1988b]. <<

[180] Muchos, muchos musulmanes están de acuerdo, y yo debo no sólo escucharles, sino hacer lo que se pueda por protegerlos y apoyarlos, en su valiente intento por reformar, desde dentro, la tradición que ellos quieren convertir en algo mejor, algo éticamente defendible. *Este* es —o mejor, quisiera ser— el mensaje del multiculturalismo, no patrocinar sutilmente la hipertolerancia racista que «respeta» doctrinas viciosas e ignorantes cuando son propuestas por representantes de estados y religiones de estados no europeos. Se puede comenzar por recomendar el libro titulado *For Rushdie* (Braziller 1994), una colección de ensayos de escritores árabes y musulmanes, muchos críticos de Rushdie, pero todos denunciando la atrozmente inmoral *fatwa* con la sentencia de muerte proclamada por el Ayatolá Jomeini. Rushdie (1994) ha dirigido nuestra atención a los 162 intelectuales iraníes quienes, valerosamente, han firmado una declaración en apoyo de la libertad de expresión. Permítaseme distribuir el riesgo cogiéndonos de la mano con ellos. <<

Notas traductor

[1t] En la edición española se ha incorporado un glosario de los términos técnicos utilizados, que aparece al final del texto. *(N. del T.)* <<

[2t] «*Evolution is a change from a no-howish untalkabourable all-alikeness by continuous sticktogheterations and somethingelsifications*», en el texto original, con palabras inventadas por W. James. *(N. del T.)* <<

Daniel C. Dennett

LA PELIGROSA IDEA DE DARWIN





